

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 01:17:11 ; Search time 12208 Seconds
(without alignments)
11919.563 Million cell updates/sec

Title: US-09-923-515-3_COPY1_5000
Perfect score: 5000
Sequence: 1 ctggagttggagacactt.....caagcttgcatcatatgac 5000

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: GenEmbl1:*
2: gb_ba:*
3: gb_hgt:*
4: gb_in:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_pro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hgt_hum:*
31: em_hgt_inv:*
32: em_hgt_other:*
33: em_hgt_mus:*
34: em_hgt_pln:*
35: em_hgt_rod:*
36: em_hgt_mam:*
37: em_hgt_vrt:*
38: em_sy:*
39: em_hgtgo_hum:*
40: em_hgtgo_mus:*
41: em_hgtgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5000	100.0	13938	9	HSALIPOA	X06290 Human mRNA
2	2687.8	54.0	4510	9	MACAPOA	J04635 Rhesus monk
3	955.4	19.1	2041	9	HSU19517	U19517 Human (apoa
4	826	16.5	1712	9	HSU19518	U19518 Human (apoa
5	442.8	8.9	1649	9	AF029691	AF029691 Papio ham
6	396.6	7.9	435	9	S79621	S79621 apolloprot
7	277	5.5	194321	2	AC027197	AC027197 Homo sapi
8	242.4	4.8	2711	9	MACCEPSC	J04697 Rhesus monk
9	236.6	4.7	1784	9	HUMPMGM	K02922 Human plasm
10	236.6	4.7	2433	6	AX463622	AX463622 Sequence
11	236.6	4.7	2497	6	AR037325	AR037325 Sequence
12	236.6	4.7	2497	6	AR082437	AR082437 Sequence
13	236.6	4.7	2497	6	AR085163	AR085163 Sequence
14	236.6	4.7	2497	6	AX448883	AX448883 Sequence
15	236.6	4.7	2497	6	HUMPLASM	M74220 Human plasm
16	236.6	4.7	2679	6	I06212	I06212 Sequence 14
17	236.6	4.7	2732	9	HSPWGR	X05199 Human mRNA
18	236.6	4.7	2753	6	A22096	A22096 plasmidogen
19	236.6	4.7	2753	6	I45623	I45623 Sequence 1
20	235	4.7	2286	6	AR105749	AR105749 Sequence
21	235	4.7	2286	6	I14104	I14104 Sequence 18
22	216.6	4.3	1134	6	AX417468	AX417468 Sequence
23	216.2	4.3	1134	6	AX395661	AX395661 Sequence
24	207.6	4.2	2771	10	BC014773	BC014773 Mus muscu
25	206	4.1	2720	10	MUSP1GN	J04766 Mouse plasm
26	202.8	4.1	194321	2	AC027197	AC027197 Homo sapi
27	199	4.0	2770	4	BTPLASMIN	X79402 B. laurus mr
28	195.6	3.9	1092	6	AX417450	AX417450 Sequence
29	190.2	3.8	1077	4	AY069985	AY069985 Canis fam
30	187.2	3.7	2745	4	EU03171	U3171 E. coli
31	182.6	3.7	2737	10	RNO242649	AJ242649 Rattus no
32	182	3.6	98357	2	AL391871	AL391871 Homo sapi
33	182	3.6	135751	2	AL396089	AL396089 Human DNA
34	182	3.6	167031	2	AC040907	AC040907 Homo sapi
35	181.6	3.6	2832	4	AF012297	AF012297 Macropus
36	172.4	3.4	141287	2	AC093404	AC093404 Pan trogl
37	172.4	3.4	141287	2	AC093404	AC093404 Pan trogl
38	169.4	3.4	5533	9	HUMAPPAKIV	L14005 Homo sapien
39	166	3.3	171363	2	AC084863	AC084863 Papio cyn
40	165.4	3.3	964	11	HSB1D8T	AL591278 H. sapiens
41	165.4	3.3	163599	9	HSU81D8T	AL109933 Human DNA
42	165	3.3	220	9	HUMPGRRRA	M86879 H. sapiens p
43	164.4	3.3	179528	2	AC084862	AC084862 Papio cyn
44	162	3.2	220	9	HUMPGRRRB	M86880 H. sapiens p
45	157.2	3.1	219	9	HUMPGRRRC	M86881 H. sapiens p

ALIGNMENTS

RESULT 1
LOCUS HSALIPOA 13938 bp mRNA linear PRI 30-MAR-1995
DEFINITION Human mRNA for apolipoprotein(a).
ACCESSION X06290 M17399 X06696
VERSION X06290.1 GI:28619
KEYWORDS apolipoprotein; apolipoprotein A; glycoprotein; lipoprotein;
plasmidogen; serine protease.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 13938)
AUTHORS McLean,T.W.
TITLE Direct Submission

QY	361	TCGACCCCAAGAGGAGCTGCGCGTCCGCGCTCCGACTGTTACCCGGGTTCCAAAGCTTAG	420
Db	361	TCGACGCCAAGAGGAGCTGCGTCCGCGCTCCGACTGTTACCCGGGTTCCAAAGCTTAG	420
QY	421	GCTCCTTCCGAACAAGACCGACTGAGCAAAAGGCCGGGGTGCAGAGAGCGTACCATG	480
Db	421	GCTCCTTCCGAACAAGACCGACTGAGCAAAAGGCCGGGGTGCAGAGAGTGTACCATG	480
QY	481	AATGGACAGAGTTATCGAGGGCACATCTCCACACTGTTCACGGAAGAACTGGCCAACT	540
Db	481	AATGACAGAGTTATCGAGGGCACATCTCCACACTGTTCACGGAAGAACTGGCCAACT	540
QY	541	TGGTCATCTTGGACACCACTCGCATGTAGTGGACCCCGCAATTAATACCCAAATGCTGC	600
Db	541	TGGTCATCTTGGACACCACTCGCATGTAGTGGACCCCGCAATTAATACCCAAATGCTGC	600
QY	601	TTGATCATGAACACTGACAGAAATCCAGATGCTGTGGCAGCTCTTATTTGTTATACGAG	660
Db	601	TTGATCATGAACACTGACAGAAATCCAGATGCTGTGGCAGCTCTTATTTGTTATACGAG	660
QY	661	GATCCCGGTGTCAAGTGTGGAGTACTGTCAACCTGACGCAATGCTCAGACGAGAGGACT	720
Db	661	GATCCCGGTGTCAAGTGTGGAGTACTGTCAACCTGACGCAATGCTCAGACGAGAGGACT	720
QY	721	GCGTCGCGCTCCGACTGTTACCCGGTTCGAAGCTGTAGAGCTTCGTCGGAACAAGCA	780
Db	721	GCGTCGCGCTCCGACTGTTACCCGGTTCGAAGCTGTAGAGCTTCGTCGGAACAAGCA	780
QY	781	CCGACTGAGCAAAAGGCTGGGGTGCAGAGAGTCTACATGTTATGACAGAGTTATGA	840
Db	781	CCGACTGAGCAAAAGGCTGGGGTGCAGAGAGTCTACATGTTATGACAGAGTTATGA	840
QY	841	GGCAGATATCTCCACACTGTGACAGGAABAACCTGCCAAGCTTGGTCAATATGACACA	900
Db	841	GGCAGATATCTCCACACTGTGACAGGAABAACCTGCCAAGCTTGGTCAATATGACACA	900
QY	901	CACCTGGCATAGTGGAGCCCGCAATACTACCCAAATGCTGGCTTATATGAATACTATGC	960
Db	901	CACCTGGCATAGTGGAGCCCGCAATACTACCCAAATGCTGGCTTATATGAATACTATGC	960
QY	961	AGGAATCCAGATGCTGGAGAGTCTCTTATTTGTTATACAGAGGATCCCGGTGCAGTGG	1020
Db	961	AGGAATCCAGATGCTGGAGAGTCTCTTATTTGTTATACAGAGGATCCCGGTGCAGTGG	1020
QY	1021	GAGTACTGCAACTGAGCGCAATGCTCAGACGAGAGAGGAGTGCCTGGCCCTCCGACT	1080
Db	1021	GAGTACTGCAACTGAGCGCAATGCTCAGACGAGAGAGGAGTGCCTGGCCCTCCGACT	1080
QY	1081	GTTACCCCGGTTCCAGGCTTAGAGGCTCCTTCCGAACAAGCACCGAGTGTAGCAAGGCT	1140
Db	1081	GTTACCCCGGTTCCAGGCTTAGAGGCTCCTTCCGAACAAGCACCGAGTGTAGCAAGGCT	1140
QY	1141	GGGGTTCAGAGAGTGTACCTATGGTATGACAGAGTTATCGGGAGCATCTCCACACT	1200
Db	1141	GGGGTTCAGAGAGTGTACCTATGGTATGACAGAGTTATCGGGAGCATCTCCACACT	1200
QY	1201	GTCACAGGAAGAACTGCGCAAGCTTGGTATATGACACCACTGCTCATATGTGAGCC	1260
Db	1201	GTCACAGGAAGAACTGCGCAAGCTTGGTATATGACACCACTGCTCATATGTGAGCC	1260
QY	1261	CCGAAATACATCCCAATGCTGGCTTGATCATGAACACTGTGAGGAATCCAGATGTGTG	1320
Db	1261	CCGAAATACATCCCAATGCTGGCTTGATCATGAACACTGTGAGGAATCCAGATGTGTG	1320
QY	1321	GCAGCTCTTATTTGTTATACGAGAGTATCCCGGTGTACAGTGGGATATCTGCAACTGAG	1380
Db	1321	GCAGCTCTTATTTGTTATACGAGAGTATCCCGGTGTACAGTGGGATATCTGCAACTGAG	1380
QY	1381	CAATGCTCAGACGAGAGGAGTCCGCTGCGGCTCCGACTGTTACCCGGGTTCCAAG	1440
Db	1381	CAATGCTCAGACGAGAGGAGTCCGCTGCGGCTCCGACTGTTACCCGGGTTCCAAG	1440

QY	1441	CTAGAGGCTCTTCCGAACAAGCACCGACTGTAGCAAAAGGCTTGAGAGTGTAC	1500
Db	1441	CTAGAGGCTCTTCCGAACAAGCACCGACTGTAGCAAAAGGCTTGAGAGTGTAC	1500
QY	1501	CATGGTATGACAGAGTTTATCGGGACACTTCCACACTGTCAACAGAAACTGCG	1560
Db	1501	CATGGTATGACAGAGTTTATCGGGACACTTCCACACTGTCAACAGAAACTGCG	1560
QY	1561	CAAGCTGTGTCATTATGACACGACACTCCATATGTGGAGCCCGAATAATCTCCAAAT	1620
Db	1561	CAAGCTGTGTCATTATGACACGACACTCCATATGTGGAGCCCGAATAATCTCCAAAT	1620
QY	1621	GCTGGCTGTATCATGAACTACTGTGAGGATTCAGATGTGTGGACGTCTTATTTAT	1680
Db	1621	GCTGGCTGTATCATGAACTACTGTGAGGATTCAGATGTGTGGACGTCTTATTTAT	1680
QY	1681	ACGAGGGATCCCGGTGTCAGGTGGGAGTACTGCACTGACGATAGCATGCTCAAGCAGAA	1740
Db	1681	ACGAGGGATCCCGGTGTCAGGTGGGAGTACTGCACTGACGATAGCATGCTCAAGCAGAA	1740
QY	1741	GGGATGCGCTTCGGGCGCTCCGAGCTGTACCCCGTTCCAAAGCTATAGAGGCTCTTCGAA	1800
Db	1741	GGGATGCGCTTCGGGCGCTCCGAGCTGTACCCCGTTCCAAAGCTATAGAGGCTCTTCGAA	1800
QY	1801	CAAGCACCAGTGTAGCAAAAGGCTGGGGTGCAGAGTGTCTACCATGTATATGACAGAGT	1860
Db	1801	CAAGCACCAGTGTAGCAAAAGGCTGGGGTGCAGAGTGTCTACCATGTATATGACAGAGT	1860
QY	1861	TATGAGGACATATCTCCACACTGTGTACAGGAAGAACTGCCAAGCTTGGTCACTATG	1920
Db	1861	TATGAGGACATATCTCCACACTGTGTACAGGAAGAACTGCCAAGCTTGGTCACTATG	1920
QY	1921	ACACCACACTGCAATGTCGGAGCCCGAGATTAATACCAAATGTGCGCTTCATCTATGAC	1980
Db	1921	ACACCACACTGCAATGTCGGAGCCCGAGATTAATACCAAATGTGCGCTTCATCTATGAC	1980
QY	1981	TACTGCAGAAATCCAGATGTGCTGGGACGTCTTATTTATATGAGGAGTCCCGGTGC	2040
Db	1981	TACTGCAGAAATCCAGATGTGCTGGGACGTCTTATTTATATGAGGAGTCCCGGTGC	2040
QY	2041	AGGTGGAGTACTCAACTGACACTGTGGACGTCTTATTTATATGAGGAGTCCCGGTGC	2100
Db	2041	AGGTGGAGTACTCAACTGACACTGTGGACGTCTTATTTATATGAGGAGTCCCGGTGC	2100
QY	2101	CCGACTGTTACCCCGGTTCCAAAGGCTTAGAGGCTCTTCGGACAAAGCACCGACTAGCAA	2160
Db	2101	CCGACTGTTACCCCGGTTCCAAAGGCTTAGAGGCTCTTCGGACAAAGCACCGACTAGCAA	2160
QY	2161	AGGCTGTGGGTGCAGAGTGTCTACCATGTATATGACAGAGTTATCGAGGCACATATCC	2220
Db	2161	AGGCTGTGGGTGCAGAGTGTCTACCATGTATATGACAGAGTTATCGAGGCACATATCC	2220
QY	2221	ACCACTGTCAAGGAAGAACCCTGCAAGCTTGGTCACTATGACACACACTCGCATAGT	2280
Db	2221	ACCACTGTCAAGGAAGAACCCTGCAAGCTTGGTCACTATGACACACACTCGCATAGT	2280
QY	2281	CGGACCCCAAGAAATCTACCCAAATGTGCGTTATCATGAACTACTGACAGGAATCCAAAT	2340
Db	2281	CGGACCCCAAGAAATCTACCCAAATGTGCGTTATCATGAACTACTGACAGGAATCCAAAT	2340
QY	2341	GCTGTGGCAGCTCTTATTTATATGAGAGGATCCCGGTGTGAGTGGGAGTACTGCAAC	2400
Db	2341	GCTGTGGCAGCTCTTATTTATATGAGAGGATCCCGGTGTGAGTGGGAGTACTGCAAC	2400
QY	2401	CTGACGCAATCTTAGACGCAAGAGGACTGCCGTGGCGCTTCGACTGTATACCCGGTT	2460
Db	2401	CTGACGCAATCTTAGACGCAAGAGGACTGCCGTGGCGCTTCGACTGTATACCCGGTT	2460
QY	2461	CGAAGCTTAGAGGCTCTTCGGAAGAAAGCAAGCACTGACAGAAAGGCTGGGGTGCAGAG	2520
Db	2461	CGAAGCTTAGAGGCTCTTCGGAAGAAAGCAAGCACTGACAGAAAGGCTGGGGTGCAGAG	2520
QY	2521	TGCTACCAATGTATATGACAGAGTTATTCAGAGGACATATCTCCACACACTGTACAGAGAA	2580

[illegible]

Query Match	54.0%; Score 2697.8; DB 9; Length 4510;
Best Local Similarity	85.4%; Pred. No. 0;
Matches 3027; Conservative	0; Mismatches 492; Indels 24; Gaps 1;
322 CCGGGGTCAGGTGGGAGTCACTGCAACCTGACAGCAATGCTCAGACGCGAAGGAGCTGCC	381
1 CCCAATGTCAAGTGGGAGTCACTGCAACCTGACACATCTCAGACGCGAAGGAGCTGCC	60
382 GTGCGGCTCCGACATGTTATACCCCGGTTCCAAAGCTTAAGAGGCTCTTTCCGAACACACCG	441
61 GTGCGACCTCCGAATGTCAACCCCGGTTCCAAAGCTTAAGAGGCTCTTTCCGAACACACCG	120
442 ACTGAGCAAAAGGCTGGGGTGTGAGGAGTGTACCTATGTAAATGACAAAGTTATCGAAGC	501
121 ACTGAGCAAAAGGCTGGGGTGTGAGGAGTGTACCTATGTAAATGACAAAGTTATCGAAGC	180
502 ACATATCTCCACATCTGTCAACAGAGAACCTGCCAAGTTGGTCTATGATACACACAC	561
181 ACATATCTCCACATCTGTCAACAGAGAACCTGCCAAGTTGGTCTATGATACACACAC	240
562 TCGCATATGTGGAACCCAGATATCTAACCAATGTGCTGTGATCATGATCTACGAGG	621
241 TCTCATATGTGGAACCCAGATATCTAACCAATGTGCTGTGATCATGATCTACGAGG	300
622 AATTCAGATGTGTGTGAGGAGCTCTTATGTTTATGAGGAGATCCCGGGTGTAGAGGGAG	681
301 AATTCAGATGTGTGTGAGGAGCTCTTATGTTTATGAGGAGATCCCGGGTGTAGAGGGAG	360
682 TACTGCAACCTGAGCAATGCTCAGACGCGAAGGAGTCCCGTGGCGCTCCGACTGT	741
361 TACTGCAACCTGAGCAATGCTCAGACGCGAAGGAGTCCCGTGGCGCTCCGACTGT	420
742 ACCCGGGTCCAAAGGCTTAGGGCTCTTCCGAACGACGACACTGAGCAAAAGGCTGGG	801
421 ACCCGGGTCCAAAGGCTTAGGGCTCTTCCGAACGACGACACTGAGCAAAAGGCTGGG	480
802 GTGCGAGATGCTACCATGTATGTATGAGACAGATTATCGAGGACATCTACCAACTGTC	861
481 GTGCGAGATGCTACCATGTATGTATGAGACAGATTATCGAGGACATCTACCAACTGTC	540
862 AAGAGAAAGACTGCGCAAGCTTGTCATCTATGACACACACTCCCATATGTGGAACCCA	921
541 AAGAGAAAGACTGCGCAAGCTTGTCATCTATGACACACACTCCCATATGTGGAACCCA	600
922 GAATACTACCAAAATGTGGCTTGTATCATGAATCTACTCAGAAATCCAGATGCTGTGCA	981
601 GAATACTACCAAAATGTGGCTTGTATCATGAATCTACTCAGAAATCCAGATGCTGTGCA	660
982 GCTCTTATTTTATATACGAGGAGATCCCGGTGTCAAGTGGGAGTACTGCAACCTGACGCA	1041
661 GCCCTTATTTTATATACGAGTATCCCAATGTCAAGTGGGAGTACTGCAACCTGACGCA	720
1042 TCGCTAGACGCAAGGAGCTGCGTCCGCGCTCGACATGTTACCCCGGTTCCAAAGCTA	1101
721 TCGCTAGACGCAAGGAGCTGCGTCCGCGCTCGACATGTTACCCCGGTTCCAAAGCTA	780
1102 GAGGCTCTCTTCGGAACAAGCACGACTGAGCAAAAGGCTGGGGTGTGACAGATGTATACAT	1161
781 GAGGCTCTCTTCGGAACAAGCACGACTGAGCAAAAGGCTGGGGTGTGACAGATGTATACAT	840
1162 GGTATGACAGAGTTATCGAGGCGCATATCTCACCACTGTGCACAGGAAGAACTGCCAA	1221
841 GGTATGACAGAGTTATCGAGGCGCATATCTCACCACTGTGCACAGGAAGAACTGCCAA	900

QY	1222	GCCTGGCTCATCTATGACACCACTCTGCAATATCTGGACCCCAAGATACTACCCAAATGCT	1281
Db	901	GCCTGGCTCATCTATGAAACCACTCTCAATATCGAACCCGGAAATACTTACCCAAATGCT	960
QY	1282	GGCTTGATCAATGAACTACTGACAGAAATCCAGATGCTGTGGCAGCTCTATTATGTTATACG	1341
Db	961	GGCTTGATCAAGAACTACTGACAGAAATCCAGATGCTGTGGCAGCTCTATTATGTTATACG	1020
QY	1342	AGGAGTCCGGGTCTCAGGTGEGAGTACTGCAACTGTAGCGAAATGCTCAGACGCAAGAGGG	1401
Db	1021	ATGATGCCCATGTCTCAGGTGEGAGTACTGCAACTGTAGCGCAATGCTCAGACGCAAGAGGG	1080
QY	1402	ACGTGCGTGGCGCTCGAGCTGTATTACCCGGTTCGAAGGCTATAGAGGCTCTTCGGAACAA	1461
Db	1081	ACTGCGTGGCGCTCGAGTGTACCCCGGTTCCGAAGGCTATAGAGGCTCTTCGGAACAA	1140
QY	1462	GCACCGACTGAGCAAAAGGCTTGGGGTGTCAGAGAGTGTACCATGTTATGAGACAGATTAT	1521
Db	1141	GCACCACTGAGCAAAAGGCTTGGGGTGTCAGAGAGTGTACCATGTTATGAGACAGATTAT	1200
QY	1522	CGAGGCACATATCTCCACCACTGTCAAGAGAAACCTGCGCAAGCTTGGTTCATCTATGACA	1581
Db	1201	CGAGGCACATATCTCACCACTGTGACAGAGAAACCTGCGCAAGCTTGGTTCATCTATGACA	1260
QY	1582	CCACACTCGCATATAGTGGGAGCCCGAGATACTACCCAAATGCTGGCTTATATGACAACTAC	1641
Db	1281	CCACACTCTCATATGTGGAGCCCGAGAAACTACCCAAATGCTGGCTTGTGTCAGAACTAC	1320
QY	1642	TGCAAGAAATCCAGATGCTGTGGAGAGTCTTATGTTATACGAGAGATCCCGGTGTACAG	1701
Db	1321	TGCGGAATCCAGATCTGTGGAGAGCCCTTGGTGTATACAAAGGATCCCAAGTGTACAG	1380
QY	1702	TGGAGATCTAGCAACTGAGGAGCAGTACTCTAGACGCAAGAGGAGTCCGCTGGCCCTCCG	1761
Db	1381	TGGAGATCTAGCAACTGTAGACAGATCTCTAGATGCAAGAGGAGTCTGTCTATGCTGTCCA	1440
QY	1762	ACTGTTACCCGGTTCCAAGCCTTAGAGGCTCCTTCGGAACAAGCACGACTGAGCAAGG	1821
Db	1441	AATATTATTCGGGTCCCAACCTTAGAGGCTTCTTGTGAACAAGAACTACTAGAGAAAC	1500
QY	1822	CCGTGGGTGAGAGGTGCTACATGCTATGAGGACAGAGTTATCGAGGACATATCTCCACC	1881
Db	1501	CCCGGGTACAGGAGGTCTACTACCTTATGACAGAGTTATCGAGGACATATCTCCACC	1560
QY	1882	ACTGTACAGAGAAACCTGCCCAAGCTTGGTCACTATGACACCACTGCGCATAGTCGG	1941
Db	1561	ACTGTACAGAGAAACCTGCCCAAGCTTGGTCACTATGACACCACTGCGCATAGTCGG	1620
QY	1942	ACCCCAAAATACACCAAAATGCTGGCTGTGATGATGAATACTGTGAGGAATCCAGATGCT	2001
Db	1621	ACCCCAAAATACATCCAAATGCTGGCTGTGATGAGGAATACTGTGAGGAATCCAGATGCT	1680
QY	2002	GTGGCAGCTCTTATTGTTTATACGAGGATCCCGGTGTCAGGTGGGAGTACTGCAACTG	2061
Db	1681	GAGATTGCGCCCTGTGTTTATACCATGAGTCCAGGTGTCAGGTGGGAGTACTGCAACTG	1740
QY	2062	ACGGAATGCTCAACGCAAGAGGAGTGGCCGTGCGCCCTCGCACTGTATACCCGGTTCOA	2121
Db	1741	ACACAAATGCTGGTGAAGAGATCAATGTGCTTGAATCTCTCACATGTGTCCCGATATCA	1800
QY	2122	AGCCTTAGAGCTCTTCCGCAACAGCACGACCTGAGCAAGGCTCGGGGTGCGAGGATGTC	2181
Db	1801	AGCACACAGGCTCTTGTGAAGAGCAACCAAGGAGCAAAAGTCCGAGAGTCCAGGATGCTC	1860
QY	2182	TACCATGTTATGAGCAGAGTTATCGAGGCACATACTCCACACTGTGTCACAGGAAGAACC	2241
Db	1861	TACCATGTTATGAGCAGAGTTATCGAGGCTCAATCTCCACACATGTGTCACAGGAAGACA	1920
QY	2242	TGCGAACCTGTGATCTATGACACCAACTGCGCAATGTGCGAGCCCGCAATACTATCCA	2301
Db	1921	TGTCAAGTCTTGCTCTTATGACACCACTGCGCACTGAGGAGCAACAGATATTATATCA	1980
QY	2302	AATCTGTGGCTTGATCAGAACTACTCGAGGAATCCAGATGCTGTGGAGCTCTTATTGCT	2361

D	b	1981	GATGGTGGCTBACCAGGAACTACTGACGAAGATCCAGATGCTGAGATTCCGGCTTGGTGT	2040
O	y	2362	TATACGAGGAGTCCGGGTGTCAGGTGGAGATCTCAACTGACGAATGCTCAGACGCA	2421
D	b	2041	TATACCATGATCCAGTGTGAGGTGGAGATCTCAACTGACGAATGCTCAGTGCACA	2100
O	y	2422	GAAGGAGCTCCCGTGGCCCTCCGACTGTATACCCCGGTTCGAACGCTAGAGGCACTCC	2481
D	b	2101	GAATCAAGTGTCTTCGCAACGTCCA-----TGGCTGTTC	2136
O	y	2482	GAACAAGCACCCAGTACGAAGAAGGCGTGGGTGACAGAGTCTACATGATATATGACAG	2541
D	b	2137	GAACAAGACCAATGAGAGAAAGCCCGGGGTCCAGAGCTGTACCATGATGATGAGCAG	2196
O	y	2542	AGTTATCGAGGCATCTACTCCACCCTGTACAGGAAACACTGCCAAGCTTGGTCACT	2601
D	b	2197	AGTTATCGAGGTCTATCTCCACCAGTGTCAAGGAAGATATGCACTGTGGTCTCT	2256
O	y	2602	ATGACACCCACATCTGCATATGTGGACCCAGAAATCTACCCAAATGTGCGCTTATCATG	2661
D	b	2257	ATGACACCCACATCTGCGATCAGAGAACATATGAAATCTACCCAAATGTGGCTGACCAAG	2316
O	y	2662	AACTACTCGAGGAATCCAGATGCTGTGGCACTCTTATTTGTTATAGAGGATCCGGT	2721
D	b	2317	AACTACTCGAGGAATCCAGATGCTGATGATTCGCCCTTGTGTTATACCATGATCCAGA	2376
O	y	2722	GTACAGTGGGAATCTGCAACCTGACGGAATGCTCAGACGGAGGAAGGACTGCGGTGCG	2781
D	b	2377	GTACAGTGGGAATCTGCAACCTGACCAATGTGTGTGATGAGAAATCAAGTGTCTTGCA	2436
O	y	2782	CCCTGACATGTTACCCGGTTCCAAGCCTTCAAGGCTCTTCGAAACAGCACGACTGAG	2841
D	b	2437	ACTCCATGTTGTCCTCCAGTCTCCAGACAGAGAGTTCCTTGTGAAGAAGCACCACTGAA	2496
O	y	2842	CAAAGGCTTGGGGTGCAGAGAGTGTCAACATGTTATGACAGAGTTATCGAGGCAATAC	2901
D	b	2497	AACAGCCCTGGGGTCCAGAGACTGTCAACAAGGTATGAGACAGATTATCGAGGCAATTC	2556
O	y	2902	TCCACCACTGTACAGAGAAGACCTGCCAAGTTGTGATGTATGACACCACTCGCAT	2961
D	b	2557	TCCACCACTATACAGGAAGAACATGTCACTGTGTTGTGTATGACCAACATCGGAT	2616
O	y	2962	AGTCGACCCCAATACTACCCAAATGCTGTGCTTGATCATGAACTACTGACGAATCCA	3021
D	b	2617	CGAGAGTCCCAATACGCTATCCAATGCTGCTGACGACGGAATATTTGCAAAATCGA	2676
O	y	3022	GATGCTGTGGACGCTTATATGTTATATACGAGAGATCCCGGTGTCAGTGGAGTACTGC	3081
D	b	2677	GATGCTGAAATTCGCCCTTGGTGTATACACATGAAATCCAGTGTAGGTGGGAGTACGC	2736
O	y	3082	AACCTGACCAATGCTCAGACGACAGAGGACTGCCGTGCGGCTCCGACTTTACCCCG	3141
D	b	2737	AACCTGACCAATGCTCAGAGTGAACAATCAATGTCTCCACAACTCCACGAGTGGTCCG	2796
O	y	3142	GTTCCAAGCTTGAAGGCTCTTCCGAACAGACCGCATGAGACAAAGGCTGGGGTGCAG	3201
D	b	2797	GTTCCAAGACAGAGGCTCTTCTGAAACAGACCACTGAAAAAGCCCTGTGTTCCAG	2856
O	y	3202	GAGTGTACCAATGTTATGACAGAGTTATCGAGGACATCTCCACATCTGCATCAGGA	3261
D	b	2857	GATGTGTACCAATGTTATGAGACAGAGTTATCGAGGACATCTCTCCACCACTGCACAGGA	2916
O	y	3262	AGAACTGCGCAAGCTTGTGATCTATGACACACACTCGGAATGTGGAGCCCGAATATC	3321
D	b	2917	AGAACTGTCAAGCTTGTGTCACTATGATGATCCACACTGGCAATCAGAGAACCCGAAATAC	2976
O	y	3322	TACCCAAATGTGCGTTGATCTATGAACTACTGACGAATCCAGATGCTGTGGCACTCT	3381
D	b	2977	TACCCAAATGTGCGCTTGACAGGAACATATCTCAGGAATTCAGATTTCTGGGAAACACCC	3036
O	y	3382	TATGTTATAGAGGATCCGGTGTCAAGTGTGAGTACTGCAACCTACGCAATGCTCA	3441

Db 3037 TGGTGTACACGACGATCATGTGTGAGTGGAGTACTGCAACCTGACAAATATCTCA 3096
QY 3442 GAGCGAAGAGGAGCCCTCCGCCCTCCGACTGTACCCCGGTGCCAAGCTAGAGCT 3501
Db 3097 GAAACGAAATCAGGTGCTAGAGACTCCCACTGTTGTCGGTTCAGGATGGAAGCT 3156
QY 3502 CCTCCGACAAAGACCCGCTAGCAAAAGGCTGGGTCAGAGAGTGTACCATGTAT 3561
Db 3157 CATTCGAAGGACGACCAACTAGCAAAACCCCTGTGTCCAGCATGTCTACCATGTAT 3216
QY 3562 GAGCAGAGTTATGAGGACCACTACCTGACACAGTGTACAGAAAGCACTGCCAAGCTTGG 3621
Db 3217 GGACAGAGTTATGAGGACCACTACCTGACACAGTGTACAGAAAGCACTGTCAATCTTGG 3276
QY 3622 TCATCATGACACCACTGCAATGAGTGGAGCCCAAGATCTACCAATGCTGGCTTG 3681
Db 3277 TCATCATGACACCACTGCAATGAGTGGAGCCCAAGATCTACCAATGAGTACTTG 3336
QY 3682 ATCATGAACTACTGAGGAAATCAGATGCTGTGCAAGCTCTTATTTATACAGAGGAT 3741
Db 3337 ACAATGAACTACTGAGGAAATCAGATGCTGTGCAAGGCCCCCTGTTTATACATGAGC 3396
QY 3742 CCGGCTGTAGGTGGAGAGTGTGCAACTGAGGCAATGCTGTGAGGAGGAGGACTCC 3801
Db 3397 CCGAGGCTGAGGCGGAGTGTGCAACTGAGGCAATGCTGTGAGGAGGAGGACTGTG 3456
QY 3802 GTGCGCCTCCGAGTGTGACCCCGGTTCCAAAGCTAGAGGCTCTTCCGAAACAAAGCAG 3861
Db 3457 GTACACCTCCGAGTGTGACCCCGGTTCCAAAGCTAGAGGCTCTTCCGAAACAAAGCAG 3516
QY 3862 ACT 3864
Db 3517 TCT 3519

RESULT 3
HS019517
LOCUS HS019517 2041 bp mRNA linear PRI 26-JAN-1996
DEFINITION Human (apoa) long mRNA, complete cds.
ACCESSION U19517
VERSION U19517.1 GI:642943
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2041)
BYRNE,C.D., SCHWARTZ,K. and LAWN,R.M.
Loss of a splice donor site at a 'skipped exon' in a gene
homologous to apolipoprotein(a) leads to an mRNA encoding a protein
consisting of a single kringle domain
Arterioscler. Thromb. Vasc. Biol. 15 (1), 65-70 (1995)
JOURNAL
MEDLINE
PUBMED 95268939
7749817
REFERENCE
LAWN,R.M.
2 (bases 1 to 2041)
Direct Submission
Submitted (05-JAN-1995) Richard M. Lawn, Stanford University School
of Medicine, Stanford University Medical Center, Falk
Cardiovascular Research Center CV 267, 300 Pasteur Drive, Stanford,
CA 94305-5246, USA

FEATURES
source
1. 2041
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6q27-ter"
/tissue_type="liver"
1. 2041
/gene="apoa"gc"
118..516
/gene="apoa"gc"
/standard_name="apolipoprotein (a) related gene C"

misc_feature
/codon_start=1
/protein_id="AA85692.1"
/db_xref="GI:642944"
/translation="MEHKEVILLILLFLKSAFTETGVSVECYHNSGYSYGTFTTV
TGRTCQAWMSMTPHOHSRTPEKYPNDGLISNCRNPDGASGFWCTYDNPVMEYCNL
TRCSDEGTVEFVPLIVIPVPSLEDSFIQVA"
1504
/gene="apoa"gc"
/note="site of the alternatively spliced variant, see
GenBank Accession number U19518"

BASE COUNT 565 a 525 c 489 g 462 t
ORIGIN
Query Match 19.1%; Score 955.4; DB 9; Length 2041;
Best Local Similarity 75.7%; Pred. No. 4.5e-286;
Matches 1367; Conservative 0; Mismatches 246; Indels 194; Gaps 7;
QY 1 CTGGGATTGGAGACACTTCTGAGACCTGCTGGCCAGTCCCAAAATGGAATAGAA 60
Db 73 CTGGGATTGGAGACACTTCTGAGACCTGCTGGCCAGTCCCAAAATGGAATAGAA 132
QY 61 GTGGTCTTCTACTCTCTTATTTTGAATTCAGACACCTGAGCAAAAGCATGTGTC 120
Db 133 GTGGTCTTCTACTCTCTTATTTTGAATTCAGACACCTGAGCAAGGCTCTCTGTG 192
QY 121 CAGATTGCTACCATGATGATGAGAGATTTATGAGGACGACTCCACACTGTGACA 180
Db 193 CAGAGAGTCTACCATGATGATGAGAGATTTATGAGGACGACTCCACACTGTGACA 252
QY 181 GGAAGACCTGCCAAGCTTGTGATGATGAGACACCATCAATCAATATAGACAGAA 240
Db 253 GGAAGACCTGCCAAGCTTGTGATGATGAGACACCATCAATCAATATAGACAGAA 312
QY 241 AACTACCAAAATGCTGCTGATCATGAACTACTGAGAGATTCAGATGCTGTGGAGCT 300
Db 313 AAGTACCAAAATGCTGCTGATCATGAACTACTGAGAGATTCAGATGCTGTGGAGCT 372
QY 301 CCTATATTTATAGAGAGATCCCGGCTGTCAGTGGAGTACTGCAACTGACGCAATGC 360
Db 373 CCTTGTGTATATAGAGAGATCCCGGCTGTCAGTGGAGTACTGCAACTGACGCAATGC 432
QY 361 TCAGACGAGAGAGGAGTCCCGGCTGCCAGCTGTATACCCGGTTCACAGCTTAGAG 420
Db 433 TCAGACGAGAGAGGAGTCCCGGCTGCCAGCTGTATACCCGGTTCACAGCTTAGAG 492
QY 421 GCTCTCTCCAGACAGACACCGACTGAGCAAAAGGCTGGAGAGTGTACATGTGT 480
Db 493 GATTCATTCTATACAG----- 508
QY 481 AATGACAGAGTATATCGAGGACATATCTCCACCACTGTACAGAGAAACCTGCCAAGCT 540
Db 509 ----- 508
QY 541 TGTGATCTATGACACCACTGCAATAGTGGACCCAGAAATACTACCAAAATGCTGCC 600
Db 509 -----TGGC 512
QY 601 TTGATCATGAACTACTGAGAAATCCAGATGCTGTGGAGCTCTTATTTATATAGAG 660
Db 513 TTGATCTCTCAATCTGAGAAATCCGATTTGTGGAGGCTCTTGTGTATATAGAG 572
QY 661 GATCCCGGTGTAGAGTGGAGTACTGCAACTGACCAATGCTTCAGACAGAGAGGACT 720
Db 573 GATCCCAAGTGTAGAGTGGAGTACTGCAACTGACAGATGCTTCACAAAGATAGGCT 632
QY 721 GCCGTGCGGCTCCGACTGTACCCCGGTTCAGAGCTTGAGGCTCTTCCGACAGCA 780
Db 633 GTGGCCCGGCTCCGACTGTATATATATATATATATATATATATATATATATATAT 692
QY 781 CCGACTGAGCAAGAGGCTGGGAGTGTGAGAGTGTACATGATGATGAGAGATATGCA 840
Db 693 CTGATTCATCAAGGCTTCAATGTAGAGTGTACATGAGAAATGAGAGATATATGCA 752

LOCUS	DEFINITION	Human (apofa19C) short mRNA, complete cds.	1712 bp	mRNA	linear	PRI 26-JAN-1996
QY	841	GGCAGATCTACTCCACCACTGTGTACAGGAAGAACCTGCCAAGTGTGTATCTATGACACA	900			
Db	753	GGCAGATCTACTCCACCACTGTGTACAGGAAGAACCTGCCAAGTGTGTATCTATGACACA	812			
QY	901	CACATCGATAGTCCGACCCCGAATACTACCCAAATGCTGGCTTGATCATCACTACTGC	960			
Db	813	CACACGACCATAGAGCTCTCGAAAGATACCCAAATGCTGACTTGATCATCACTACTGC	872			
QY	961	AGGAATCCAGATGCTGTGGCAGCTCTTATTTGTTATACAGAGGATCCCGTGTGACGTGG	1020			
Db	873	AGGAATCCAGATGCTGTGGCAGCTCTTATTTGTTATACAGAGGATCCCGTGTGACGTGG	931			
QY	1021	GAGATCTGCAACCTGACGCAATGCTCAGACAGCCAGAGGAGTGGCGCTGCCCTCCGACT	1080			
Db	932	GAGATCTGCAACCTGACGCAATGCTCAGACAGCCAGAGGAGTGGCGCTGCCCTCCGACT	991			
QY	1081	GTTATCCCCGGTTCGAAGCTTAGAGGCTCTTCCGAACAGCAGCCGACTGAGCCAAAGGCTT	1140			
Db	992	GTTATCCCCGGTTCGAAGCTTAGAGGATCTTCCGAACAGCAGCCGACTGAGCCAAAGGCTT	1051			
QY	1141	GGGGGACGAGAGTGTCTACCATGTTAATGGAGAGATTTCGAGGACATCTACCAACCT	1200			
Db	1052	TGGGAGCAGCAGTGTCTACCATGTTAATGGAGAGATTTCGAGGACATCTACCAACCT	1111			
QY	1201	GTCACAGGAGAACCTGCGCAAGCTTGTGTCATCTATGACACACACTCGCATATGTGGACC	1260			
Db	1112	GTCACAGGAGAACCTGCGCAAGCTTGTGTCATCTATGACACACACTCGCATATGTGGACC	1163			
QY	1261	CCAGATATCTACCCAAATGCGGCTGATCTATGAACTCTCCAGGATATCCAGATGCTGTG	1320			
Db	1164	CCAGATATCTACCCAAATGCGGCTGATCTATGAACTCTCCAGGATATCCAGATGCTGTG	1202			
QY	1321	GCAGCTCTTATTTGTTATATGAGGAGATCCCGGTGTGAGGTGGAGTACTGCAACCTGACG	1380			
Db	1203	AAACAAACCTGTGTGTATCACAACCTGATCGTGTGAGGTGGAGTACTGCAACCTGACA	1262			
QY	1381	CAATGCTCAGACGAGAAAGGAGTCCGTCCGCTCCGACTGTACCCCGGTTCCATGC	1440			
Db	1263	CAATGCTCAGAAACAGAA-TCAGTGTCTTAAGACTCCCACTGTGTTCCCGTCCAAAGC	1321			
QY	1441	CTAGAGGCTCTTCCGAACAAGCACACCGACTGAGCAAAAGGCTGGGTGCAGAGTGTAC	1500			
Db	1322	ATGAGAGCTCTATTTGGAAGCAGACACCACTAGCAAAACCCCTGTGTCTCCGCGAGTGTAC	1381			
QY	1501	CATGTTAATGACAGAGTTATTCAGAGCACAATCTCACCACTGTCTACAGGAAGAACTGTC	1560			
Db	1382	CATGTTAATGACAGAGTTATTCAGAGCACAATCTCACCACTGTCTACAGGAAGAACTGTC	1441			
QY	1561	CAAGCTTGTGATCATATGACACACACATCG-CATAGTGGAGCCCAAGATATCTACCAAA	1619			
Db	1442	CAATTTTGTGTCATCATATACACACACCCGGTTCATCAGAGAACCCCAAGAAATACCCAAA	1501			
QY	1620	TGCTGGCTTGTATCATGAACACTACTGACGGAATTCAGAGTCTGTGGCAGCTCCTTATGTTA	1679			
Db	1502	TGATGGCTCTGACAAATGAACTACTGACGGAATTCAGATCCGATACAGGCCCTTGGTGT	1561			
QY	1680	TACGAGGATATCCCGGTGTGCAAGTGGGAGTACTGCAAACTGACGCAATCTCAGAGCGTCA	1739			
Db	1562	TACCAACAACCCCGACGTCAGAGTGGGAGTACTGCAAACTGACGCGATCTCAGA--CAGA	1619			
QY	1740	AGGAGTCCGCTCGCGCTCCGACTGTATCCCGGTTTCCAAAGCCCTAGAGGCTCTCTTCCGA	1799			
Db	1620	AGGAGTCCGCTCGCGCTCCGACTGTATCCCGGTTTCCAAAGCCCTAGAGGCTCTCTTCCGA	1679			
QY	1800	ACAACTA 1806				
Db	1680	ACAACTA 1686				

ACCESSION	U19518
VERSION	U19518.1
KEYWORDS	GI:642945
SOURCE	.
ORGANISM	Homo sapiens.
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases 1 to 1712)
AUTHORS	Byrne,C.D., Schwartz,K. and Lawn,R.M.
TITLE	Loss of a splice donor site at a 'skipped exon' in a gene homologous to apolipoprotein(a) leads to an mRNA encoding a protein consisting of a single kringle domain
JOURNAL	Arterioscler. Thromb. Vasc. Biol. 15 (1), 65-70 (1995)
MEDLINE	95268939
PUBMED	7749817
REFERENCE	2 (bases 1 to 1712)
AUTHORS	Lawn,R.M.
TITLE	Direct Submission
JOURNAL	Submitted (05-JAN-1995) Richard M. Lawn, Stanford University School of Medicine, Stanford University Medical Center, Falk Cardiovascular Research Center CV 267, 300 Pasteur Drive, Stanford, CA 94305-5246, USA
FEATURES	location/Qualifiers
Source	1..1712
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="6"
	/map="6q27-ter"
	/tissue_type="liver"
	1..1712
gene	/gene="apoargC"
	118..516
CDS	/gene="apoargC"
	/standard_name="apolipoprotein (a) related gene C"
	/codon_start=1
	/protein_id="AA85693.1"
	/db_xref="GI:642946"
	/translation="MEKEVYLILFLKSNPEWEGYNSQCYHSNGSYRGYFTTV TGRCDQMSSSTPHQHRSIPERKIPINDGLISICRNPDSCSAGPMCTITTDVPRMFCML TRCSDDEGTVPVLPIVIPSLDSEFIQVA"
	1504
misc_feature	/gene="apoargC"
	/note="site of the alternatively spliced variant, see Genbank Accession number U19517"
BASE COUNT	490 a 443 c 419 g 360 t
ORIGIN	
Query Match	16.5%; Score 826; DB 9; Length 1712;
Best Local Similarity	74.2%; Pred. No. 9e-246;
Matches 1216; Conservative	0; Mismatches 230; Indels 192; Gaps 6;
OY	1 CTGGGATTGGGACACACTTCCTCGACACTGCCTGGCGCATCCCAAAATGGAACATPAGGA 60
DB	
73	CTGGGATTGGGACACACTTTCTGGGCACTGCTGGGCCAATGCCAAATGAACATPAGGA 132
OY	61 GTGGTTCTTCACTCTTTTATTTTCTGAAATCAGCAGCACCTGAGCAAAGCATGTGTC 120
DB	
133	GTTGTTCTTCACTCTGTATTATTCTGAAGTCAGCACCGACTGAGACAGGGCCTTGTC 192
OY	121 CAGGATTGTCACATGCGATGAGACAGATTATCAGGACGACTCCACCACCTGTACA 180
DB	
193	CAGGAGTCTACCAAGTAATGAGACAGATTATCAGGACACATACCTTACCACTGTCCA 252
OY	181 GGAAGACCTGCAAGCTTGGTCATCTATGACACACCAATCAATATNAGACACAGAA 240
DB	
253	GGAAACACTCGCAAGCTTGGTCATCTATGACGCCACACACAGACAGTAGAAGAACCCAGA 312
OY	241 AACTGACCCAAAATGCTGGCTTGATCATGAACTACTGACAGAAATCCAGATGCTGTGGCACT 300
DB	
313	AAGTACCCAAATGATGGCTTGATCTCGAACTATCTGCAGAAATCCGGAATGTTCGGCAGGC 372
OY	301 CCTTATTGTTATAGCAGGAGATCCGGGTCAAGTGGGAGTACTGCAACCTCAGCAATGC 360

Db 373 CCTTGTTTATACGAGGATCCCAATGTCAAGTGGAGTACTGCAACCTGACAGCGTGC 432
QY 361 TTAGACGAGAAAGGAGTCCCTCCGCGCTCGACTGTTACCCCGGTTCCAAAGCTTAGAG 420
Db 433 TCAGACAGTGAAGGAGCTGTGTTCTGCTCTGCACTGTTATCCAGTTCCAAAGCTTAGAG 492
QY 421 GCTCCCTCCGACCAAGCAACCCACTAGAGAAAGGCTTG6GGGTCCAGAGTGTACCATGT 480
Db 493 GATTATTTATACAAAG----- 508
QY 481 AATGACAGAGTTATCGAGGACATACTCCACCCTGTCCAGAAAGAACTCCCAAGCT 540
Db 509 ----- 508
QY 541 TGGTCACTATATACACACACACTGCAATAGTCGACCCTCAGAAATACCTACCAATCTGGC 600
Db 509 -----TGGC 512
QY 601 TTGATCATGACTACTGCAAGAAATCCAGATGCTGTGGCAGCTCCTTATTTGTTATACGAG 660
Db 513 TTGATCTCCAACTACTGCAAGAAATCCGAGTTGTGGCAGGCCCTTGTTATACACAG 572
QY 661 GATCCCGGTGTGAGTGGAGTACTGCAACTGACGCAATGCTCAGACGACAGAGGAGCT 720
Db 573 GATCCCAAGTCAAGTGGAGTACTGCAACTGACAGGATGCTCAGACAAAGATAGGGCT 632
QY 721 GCCGTGGCGCTCCGACTGTTACCCGGTTCACAGCCTTAGAGGCTCTTCCAGACAGCA 780
Db 633 GTGGCGCGCTCTGACTATTAATCCGGTTCACAGACGAGAGATACCTTCCAAACAGCA 692
QY 781 CGGACTGAGCAAAAGGCTGGGGGTGAGAGTGTACTACAGTGAATGAGACAGATTATGA 840
Db 693 CTGATTTGATCCAAAGGCTTTCATATGTAGAGTGTACTACCAATGAAATGACAGATTATGA 752
QY 841 GGCACATCTCCACACTGTACAGAGAAAGAACTGCGCAAGCTTGTCTATGTAGACCA 900
Db 753 GGCACATCTCCACACTGTACAGAGAAAGAACTGCGCAAGCTTGTCTATGTAGACCA 812
QY 901 CACTCCGATAGTGGAGCCCAAGAAATCTACCCAAATGCTGGCTTATCATGAAATCTACTGC 960
Db 813 CACCACACAGTAGGACCTCAGAAAGTATCCCAAAATGCTACTGTATCATGAAATCTACTGC 872
QY 961 AGGAATCCAGATGCTGTGGAGCTCTTATTTGTTATGAGGAGTCCCGTGTGAGGCTG 1020
Db 873 AGGAATCCAGATCTCTGTGAGAG-GCCTTGTGTTTACACATGATGCCAAAGTCAGATGG 931
QY 1021 GAGTACTGCAACCTGACGCAATGCTCAGAGCAGAAAGGAGTCCGCTGCGCTCCGACT 1080
Db 932 GAGTACTGCAACCTGACAGCATGCTCAGACAGAAAGGAGTCTGAGTGGCTGTGACT 991
QY 1081 GTTACCCCGGTTCCAAAGCTTAGAGGCTCTTCCGAACAAAGCAGCAGTACGCAAAAGCCT 1140
Db 992 GTTATCCCGGTTCCAAAGCTTAGAGGATCCTTCCAAACACACACAGCAGCAAGGCT 1051
QY 1141 GGGGTGAGAGTGTACTACATGTTATGACAGAGTATGAGAGGACATCTACTCCACT 1200
Db 1052 TGGGAGACAGAGTGTACTACAGCGGTATGAGACAGATTATGAGGACACATACTTACACT 1111
QY 1201 GTACAGAGAAAGAACTGCGCAAGCTTGGTCACTATGACACACACATGCGATAGTCGGAAC 1260
Db 1112 GTACAGAGAAAGAACTGCGCAAGCTTGGTCACTATGAGCGCCACATCAGCACA----- 1163
QY 1261 CCAGAAATACACCAAAATGCTGGCTGATCATGACTACAGTACGAGGAATCCAGATGCTGTG 1320
Db 1164 -----TGGCTGACCGAGAAC-ACCTGACAGGAATCCAGATTTCTGGG 1202
QY 1321 GCAGCTCCTTATTTATATACGAGGAGTCCCGGTGTCAAGTGGAGTACTGCAACTGACG 1380
Db 1203 AAACAACCTCGTGTATACACACATCCGTGTGTGAGTGGAGTACTGCAACTGACG 1262
QY 1381 CAATGCTCAGACGAGAGGAGTGCCTGTGGCGCTCCGACTGTACCCCGGTTCCAAAGC 1440
Db 1263 CAATGCTCAGAAACAGAAATC-AGTGTCTTAGAGACTCCCACTGTGTGTTCCCGTTCCAAAGC 1321

QY 1441 CTAGAGGCTCCTTCGAAACAAGCACCGACTGAGCAAAAGCCTGGGGTGCAGAGTGTCTAC 1500
Db 1322 ATGAGAGCTACTTTCTGAAGCAGCACCAACTGAGCAAAACCCCTGTGGTCCGAGTGTCTAC 1381
QY 1501 CATGTAATGAGCAGAGTATTTCGAGGACATCTCCACCACTGTCCACAGGAGAACCTTCG 1560
Db 1382 CATGTAATGAGCAGAGTATTTCGAGGACATCTCCACCACTGTCCACAGGAGAACCTTCG 1441
QY 1561 CAAGCTTGTCTATATGACACCACTCG-CATAGTCGGAGACCCCAATATCTACCCAA 1619
Db 1442 CAATCTTGGTCACTCATGACACCAACCGGTGCATCAGAGAGACCCAGAAATACCCAA 1501
QY 1620 TGGTGGCTTGATCATGAA 1637
Db 1502 TGACTCTATGTTTGGGAA 1519

RESULT 5
AF029691 1649 bp mRNA linear PRI 28-JAN-1998
LOCUS
DEFINITION Papio hamadryas apolipoprotein a (BABAPOA) mRNA, partial cds.
ACCESSION AF029691
VERSION AF029691.1 GI:2815615
KEYWORDS
SOURCE
ORGANISM
Papio hamadryas.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Papio.
REFERENCE
AUTHORS 1 (bases 1 to 1649)
Cox,L.A., Jett,C. and Hixson,J.E.
TITLE Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice Site Mutation is Associated with Deletion of a Single Exon in a Null Allele
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1649)
Cox,L.A., Jett,C. and Hixson,J.E.
AUTHORS Direct Submission
TITLE Submitted (10-OCT-1997) Genetics, Southwest Foundation for Biomedical Research, P.O. Box 760549, San Antonio, TX 78245-0549, USA

FEATURES
source
location/Qualifiers
1..1649
/organism="Papio hamadryas"
/db_xref="taxon:9557"
<1..1649
/gene="BABAPOA"
<1..1366
/gene="BABAPOA"
/codon_start=2
/product="apolipoprotein a"
/protein_id="BAB97886.1"
/db_xref="GI:2815616"
/translation="STVTYTGNCQMSMTPEHOHSRTPKNYPNAGLTIRNYCRNPDAEI
RPMCYTMDPSVRYMEYCNLTQCLIVTESSVLETLIVPDPQSASEEPEVSDVDC
YHGDGOSRGSEFTTNGRQOSMSMTPHQKRTPEHNDGLTMYCNPADYGP
WCPETMDPSVRYMEYCNLTGRCSDTETGVYPLIVLIPSLERASQASSSFCGKDYEP
KKCRGRVYGGCVAAHSHRPMQVSLRTFRGKHFQGGTLLISEPWLTARCLEMSPPRSS
YVILGAHQEYNLDSHVOELVSKLSESPGADIALKLSRPALITDKVLPALCPSPN
YVITAMTECYLTGMEETGTGAGLLEEARLPIENTVNCNRYELNGRVASTELCSGH
LAGGTDSOQDSDGSPVVCFDKDKYILRGITSMWPGCARPNKPGYVAVSSFTVIBGV
MRNN"
2..91
/gene="BABAPOA"
/note="exon A; encodes kringle IV 9a"
92..271
/gene="BABAPOA"
/note="exon B; encodes kringle IV 9b"
272..434
/gene="BABAPOA"
/note="exon C; encodes kringle IV 10a"
435..614
exon


```

Db 301 CCTATTGTTATGAGTGGATCCCATGTGAGTGGAGTACTGCACCTGACACATGCG 360
Qy 361 TCAGACACAGAGGAGTGGCGTGGCGCTCCGACTGTTACCCGGTTCACACCTAGAG 420
Db 361 TCAGATGCAGAGGAGTGGCGTGGCGACCTCCGAAATGTCACCTGTGTCACACCTAGAG 420
Qy 421 GCTCCTTCGACACAA 435
Db 421 GCTCCTTCGACACAA 435

RESULT 7
AC027197 194321 bp DNA linear HMG 24-AUG-2002
LOCUS Homo sapiens chromosome 6 clone RP11-189014 map 6, LOW-PASS
DEFINITION SEQUENCE SAMPLING.
AC027197
VERSION AC027197.3 GI:9887745
KEYWORDS HMG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE 1 (bases 1 to 194321)
JOURNAL Homo sapiens chromosome 6, clone RP11-189014
REFERENCE
AUTHORS 2 (bases 1 to 194321)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choquel, Y., Colangelo, M., Collins, S.,
Collins, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J.S., Gage, D.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Larocque, K., Lamazares, R., Landers, T., Lehoczeky, J.,
Levine, R., Lien, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McNeeters, R.,
Medlim, J., Menus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tefaye, S., Theodore, J., Tirelli, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
JOURNAL Submitted (28-MAR-2000) Whitehead Institute/MIT Center for Genome
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS 3 (bases 1 to 194321)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Beda, F., Boguslavsky, L.,
Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
Choquel, Y., Colangelo, M., Collins, S., Collins, A., Cooke, P.,
DeRellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,
Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,
Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L.,
Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Larocque, K.,
Lamazares, R., Landers, T., Lehoczeky, J., Levine, R., Lien, C., Liu, G.,
MacDonald, P., Margulis, N., McCarthy, M., McEwan, P., McKernan, K.,
McNeeters, R., Medlim, J., Menus, L., Mihova, T., Mienga, V.,
Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,
O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K.,
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Rieback, M., Riley, R.,
Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,
Sougnuez, C., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tefaye, S., Theodore, J.,
Tirelli, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,

```

TITLE
JOURNAL

COMMENT

Zimmer, A. and Zody, M.

Direct Submission

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Aug 24, 2000 this sequence version replaced 91:979792.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: I7625

Center clone name: 189_O_14

* NOTE: This record contains 224 individual

* sequencing reads that have not been assembled into

* contigs. Runs of N are used to separate the reads

* and the order in which they appear is completely

* arbitrary. Low-pass sequence sampling is useful for

* identifying clones that may be gene-rich and allows

* overlap relationships among clones to be deduced.

* However, it should not be assumed that this clone

* will be sequenced to completion. In the event that

* the record is updated, the accession number will

* be preserved.

* 1 741: contig of 741 bp in length

* 742 841: gap of 100 bp

* 842 1594: contig of 753 bp in length

* 1595 1694: gap of 100 bp

* 1695 2442: contig of 748 bp in length

* 2443 2542: gap of 100 bp

* 2543 3274: contig of 732 bp in length

* 3275 3374: gap of 100 bp

* 3375 4104: contig of 730 bp in length

* 4105 4204: gap of 100 bp

* 4205 4929: contig of 725 bp in length

* 4930 5029: gap of 100 bp

* 5030 5737: contig of 708 bp in length

* 5738 5837: gap of 100 bp

* 5838 6574: contig of 737 bp in length

* 6575 6674: gap of 100 bp

* 6675 7382: contig of 708 bp in length

* 7383 7482: gap of 100 bp

* 7483 8227: contig of 745 bp in length

* 8228 8327: gap of 100 bp

* 8328 9072: contig of 745 bp in length

* 9073 9172: gap of 100 bp

* 9173 9920: contig of 748 bp in length

* 9921 10020: gap of 100 bp

* 10021 10740: contig of 720 bp in length

* 10741 10840: gap of 100 bp

* 10841 11566: contig of 726 bp in length

* 11567 11666: gap of 100 bp

* 11667 12407: contig of 741 bp in length

* 12408 12507: gap of 100 bp

* 12508 13354: contig of 747 bp in length

* 13355 13354: gap of 100 bp

* 13355 14095: contig of 741 bp in length

* 14096 14195: gap of 100 bp

* 14196 14933: contig of 738 bp in length

* 14934 15033: gap of 100 bp

* 15034 15755: contig of 722 bp in length

* 15756 15855: gap of 100 bp

* 15856 16584: contig of 729 bp in length

* 16585 16684: gap of 100 bp

* 16685 17433: contig of 749 bp in length

* 17434 17533: gap of 100 bp

* 17534 18286: contig of 753 bp in length

* 18287 18386: gap of 100 bp

* 18387 19127: contig of 741 bp in length

[illegible][illegible]

[illegible]

Query Match	Score	DB	Length
Best Local Similarity	58.6%	Pred. No. 3.2e-62	2433
Matches 547: Conservative	0	Mismatches 72: Gaps 5:	

BASE COUNT	678 a	607 c	616 g	532 t
ORIGIN				

FEATURES	source
CDS	<p>1..2433</p> <p>1..2433</p> <p>/note="unnamed protein product"</p> <p>/codon_start=1</p> <p>/protein_id="C6DA2143.1"</p> <p>/db_xref="GI:21886383"</p> <p>/translation="MEHREYVLLLLFLKSGGEPFLDYVNTQASLSFSTKQLGAGSIEBCANACEDEDEFTCRAPVYHSKEQCYIAENKRSIIRKMDVLEFKYVLSQCTGNGKNYRGITMSKTNGITCQKWSSTSPHRRSPATHPSEGLEENYCNPNDDQGPWCYTTPERKRYDCDILBCEBECMHCSEENDGKISITMSGLCQAMQSQSHAFGLYPSFEPNNKLNKRYCNRDPDLBECBPMCFITDPNKRMLCDIIPCTTPSSGSPYQCLKGTENYRGNAVAVYSGTCHQMSGAOLPATPNTPENFCKMLDENTYCRANDGKRAVPCHTTNSOYRMEXCKIPSCDSQSPVSTEOALATAPLELTPYQDCYHGDSQYRGTSSTTGTGKCSQMSWTFPHRHQKTPENYPAAGLIMYCNCRPDADGPMCFETDPSVREYCNLRKCSGTASVYAAPVYLIPDVTPESECDMGNGKYGKGRATYVGTGFCODMAOAEHRHSITTPENPRACLEKRYCNPDGQVGMCTYKPRKLDYCDVPOCAASPDGRAPQVERKCKGRVYVGCYAHBSRPMQVSLTRFQMHFCGTLISBPVITPAHCEKSPRSSTKVIILGAHOEVNLEPHVDIEVSRLEPEPRKDLALTKLSSPVIITDKVTPACGSPNIVYADRTGCTGTGPGTGTGKAGLLEAOPLVTEKNYCNRYEFLNGRQSPSELGAGHAGVAGTDCSGDSGGLYCFEKDKYIILAGTISWGLGCAKRPKKPVYVRSRFTWIEGVRNN"</p>

REFERENCE	AUTHORS	TITLE
JOURNAL		

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
AX463622	Sequence 1 from Patent WO0250290.	AX463622	AX463622.1	GI:21886382	human.

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
Db	686	710	791	686	710	791	686

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
Db	686	710	791	686	710	791	686

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
Db	686	710	791	686	710	791	686

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
Db	686	710	791	686	710	791	686

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
Db	686	710	791	686	710	791	686

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
Db	686	710	791	686	710	791	686

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
Db	686	710	791	686	710	791	686

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
Db	686	710	791	686	710	791	686

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
Db	686	710	791	686	710	791	686

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
Db	686	710	791	686	710	791	686

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
Db	686	710	791	686	710	791	686

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
Db	686	710	791	686	710	791	686

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
Db	686	710	791	686	710	791	686

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
Db	686	710	791	686	710	791	686

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
Db	686	710	791	686	710	791	686

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
Db	686	710	791	686	710	791	686

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
Db	686	710	791	686	710	791	686

Db	Accession	Length	Score	Expect	Ident	Mismatches	Gaps
----	-----------	--------	-------	--------	-------	------------	------

809 AGTCTACATGTAATGACAGAGTTATGAGGACATACCTCCACACTGTCCAGAGAA 868
821 AGTGTGAAGGGAACAGGGAACATATGCGGGAATGTGGTGTATCGTTTCGGGC 860
869 GAACCTCCCAAGCTTGGTCACTATGACACACACTGCTAGTTCGAGCCCAAGTAAT 928
881 ACACCTGTGACACTGAGAGTGACAGACCCCTGACACACATTAACAGACACGAAACT 940
929 ACCCAATGCTGCTTATCATCATGACTGACAGAAATCCAGATGCTGTGGACCTCTT 988
941 TCCCTCCAAAATTTGGATGAAATCTACGCGCATCTTGACGGAAGGAGCCCAT 1000
989 ATTGTATACAGAGGATCCCGGTGTCAGTGGAGTACTGCAACCTGACCAATGCTCAG 1048
1001 GGTGCTATCAACCAACGCAAGTGGGTGGAGTACTTAAGATACCT----- 1051
1049 ACGGAAAGGAGTGCCTGCGCTCGACTGTTACCCCGGTTCCAAAGCTTAGAGCTC 1108
1052 -----CCTGTGACTCTCCCGCATCCAGTATCCAG-----GAACAT 1084
1109 CTTCGGAACAGCACCCAGTACAGAAAGGCTGGGGTGCAGAGTGTACCATGTATG 1168
1085 TGGTCCACAGCACACCTGATACCCCTGTGTGTCAGAGTGTCTACCATGTATG 1144
1169 GACAGATTATCGAGGACATATCTCCACACTGTACAGAGAAAGCTGCCAAGCTGTG 1228
1145 GACAGAGTACCGAGGACATCTCTCCACACACACAGGAAAGATGTCAAGCTTGTG 1204
1229 CATCTATGACACCACTCGCATAGTGGAGCCCAAGTAATCTACCAATGTGTGCTTGA 1288
1205 CATCTATGACACCACTCGCATAGTGGAGCCCAAGTAATCTACCAATGTGTGCTTGA 1264
1289 TCATGAATCTATGAGGAATCAGATGTGTGCAAGCTCCTTATTTTATACAGGAGATC 1348
1255 CAATGAATCTATGAGGAATCAGATGTGTGCAAGCTCCTTATTTTATACAGGAGATC 1324
1349 CCGGTGTGAGTGGAGTACTGCAACCTGACGCAATGCTCAGAGCAAGAGGAGCTCG 1408
1325 CCAGGTGAGTGGAGTACTGCAACCTGAAATAATGCTGAGAAAGAGGAGTGTG 1384
1409 TCGGCTCCGACGTTATACCCGGTTCAGAGCTTCCCTCCGAAACAGACACGA 1468
1385 TAGACCTCCGCTGTGTCTGCTGCTGATGAGAGTCTTCGGAAGAA----- 1437
1469 CTGAGCAAAAGGCTTGGGTTGACGAGTGTACATGTAATGACAGAGTATTCAGAGCA 1528
1438 -----GACTGTATGTTTGGGAATGGGAAAGATACCGAGGCA 1474
1529 CATATCTCACCACGTGTACAGAAAGACCTGCAAGCTTGGTCAATGATGACACCACT 1588
1475 AGAGGGCGACCACTGTATCTGAGGAGCCATGTCAGAGTGGGCTGCCAGAGGCCATA 1534
1589 CGCATAG---TCGAGCCCAATAATACCAATGTGTGTATGATGAATACAGCA 1645
1535 GACACGAGATTTTCACTCCACAGCAAAATCCAGGGGGCTGTGAAAAAATTAATGACC 1594
1646 GGAATCCAGAA---TGCTGTGGACGCTCTTATTTATACAGAGGATCCCGGTGACGT 1702
1595 GTAACCTGATGTGATGTATGATGTGCTGTCTACACAGCAAAATTCAGAAATCTT 1654
1703 GGGAGTACTGCAACCTGACGCAATGTCTCAGAGC 1736
1655 ACGACTACTGTGATGTCTCCTCAGTGTGGGCC 1688

RESULT 11
LOCUS AR037325 2497 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 12 from patent US 5801146.
ACCESSION AR037325
VERSION AR037325.1 GI:5955181
KEYWORDS
SOURCE Unknown.

ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2497)
AUTHORS Davidson, D.J.
TITLE Compound and method for inhibiting angiogenesis
JOURNAL Patent: US 5801146-A 12 01-SEP-1998;
FEATURES
source 1. 2497
BASE COUNT 691 a 624 c 637 g 545 t
ORIGIN

Query Match 4.7%; Score 236.6; DB 6; Length 2497;
Best Local Similarity 58.6%; Pred. No. 3,2e-62;
Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;

809 AGTCTACATGTAATGACAGAGTTATGAGGACATACCTCCACACTGTCCAGAGAA 868
870 AGTGTGAAGGGAACAGGGAACATATGCGGGAATGTGGTGTATCGTTTCGGGC 929
869 GAACCTCCCAAGCTTGGTCACTATGACACACACTGCTAGTTCGAGCCCAAGTAAT 928
930 ACACCTGTGACACTGAGAGTGACAGACCCCTGACACACATTAACAGACACGAAACT 989
929 ACCCAATGCTGCTTATCATCATGACTGACAGAAATCCAGATGCTGTGGACCTCTT 988
990 TCCCTCCAAAATTTGGATGAAATCTACGCGCATCTTGACGGAAGGAGCCCAT 1049
989 ATTGTATACAGAGGATCCCGGTGTGAGTGGAGTACTGCAACCTGACCAATGTCTG 1048
1050 GGTGCTATCAACCAACGCAAGTGGGTGGAGTACTTAAGATACCT----- 1100
1049 ACGGAAAGGAGTGCCTGCGCTCGACTGTTACCCCGGTTCCAAAGCTTAGAGCTC 1108
1101 -----CCTGTGACTCTCCCGCATCCAGTATCCAG-----GAACAT 1133
1109 CTTCGGAACAGCACCCAGTACAGAAAGGCTGGGGTGCAGAGTGTACCATGTATG 1168
1134 TGGTCCACAGCACACCTGATACCCCTGTGTGTCAGAGTGTCTACCATGTATG 1193
1169 GACAGATTATCGAGGACATATCTCCACACTGTACAGAGAAAGCTGCCAAGTGTG 1228
1194 GACAGGCTACGAGGACATCTCTCACCACACACAGGAAAGATGTGAGTGTG 1253
1229 CATCTATGACACCACTCGCATAGTGGAGCCCAAGTAATCTACCAATGTGTGCTTGA 1288
1254 CATCTATGACACCACTCGCATAGTGGAGCCCAAGTAATCTACCAATGTGTGCTTGA 1313
1289 TCATGAATCTATGAGGAATCAGATGTGTGCAAGCTCCTTATTTTATACAGGAGATC 1348
1314 CAATGAATCTATGAGGAATCAGATGTGTGCAAGCTCCTTATTTTATACAGGAGATC 1373
1349 CCGGTGTGAGTGGAGTACTGCAACCTGACGCAATGCTCAGAGCAAGAGGAGCTCG 1408
1374 CCAGGTGAGTGGAGTACTGCAACCTGAAATAATGCTCAGAGAAAGAGGAGTGTG 1433
1409 TCGGCTCCGACGTTATACCCGGTTCAGAGCTTCCCTCCGAAACAGACACGA 1468
1434 TACACCTCCGCTGTGTGCTGCTGCTGATGAGAGTCTTCGGAAGAA----- 1486
1469 CTGAGCAAAAGGCTTGGGTTGACGAGTGTACATGTAATGACAGAGTATTCAGAGCA 1528
1487 -----GACTGTATGTTTGGGAATGGGAAAGATACCGAGGCA 1523
1529 CATATCTCACCACGTGTACAGAAAGACCTGCAAGCTTGGTCAATGATGACACCACT 1588
1524 AGAGGGCGACCACTGTATCTGAGGAGCCATGTCAGAGTGGGCTGCCAGAGGCCATA 1583
1589 CGCATAG---TCGAGCCCAATAATACCAATGTGTGTATGATGAATACAGCA 1645
1584 GACACGAGATTTTCACTCCACAGCAAAATCCAGGGGGCTGTGAAAAAATTAATCTGCC 1643
1646 GGAATCCAGAA---TGCTGTGGACGCTCTTATTTATACAGAGGATCCCGGTGACAGT 1702

D	b194	GACGAGCTACCGAGGCATCTCCACACACCACAGGAANAAGTGTACGCTTG6T	1253
OY	1229	CATCTATGACACCACTTCGATAGTGGGACCCAGATACTACCAAATTGCTGGCTTGA	1288
D	1254	CATCTATGACACCACTTCGAGCAGGAAGAACCCAGAAACTACCAATCTGTCGCCCTGA	1313
OY	1289	TCAATGAACTACTGCAGGAATTCATATCTGTGGGACGCTCCTATTGTATTACGAGGATC	1348
D	1314	CAATGAACTACTGCAGGAATTCATATCTGTGGGACGCTCCTATTGTATTACGAGGATC	1373
OY	1349	CCGGTGCAGGTGGAGTACTGCACAACCTGACGAATGCTGCAGCACAGAAGGAGCTGGC	1408
D	1374	CCAGCTGACGTTGGAGTAGTACTGCACAACCTGAAAAAATGCTGCAGAACAGAACGAGTGTG	1433
OY	1409	TCGCGCTCCGAGCTGTTACCCCGCTTCCAGCCTAGAGGCTCCTTCGAGACAGCACCGA	1468
D	1434	TAGCACTCCCGCTGTTGCTCCTGCTCCAGATGTAAGAGACTCCTCCGAMAAA-----	1486
OY	1469	CTGAGCAAAAGCCTGGGGTGCAGAGTGCATCCATGGATAGAGACAGAGTATGAGGCA	1528
D	1487	-----GACTGTATGTTTGGGAATGGAGAAAGATACCGAGCA	1533
OY	1529	CATACCTCACCACCTGTACAGAGAAACCTGCAGAGCTTGGTCACTATGACACCACT	1588
D	1524	AGAGGGCGACCACTGTACTGGAGCGCATCTCCAGAGACTGGCGTCCAGAGGCCCAT	1583
OY	1589	CGCATAG---TCGGAAACCCCAATATCATACCCCAATGCTGCTGTATCATGAATCACTGCA	1645
D	1584	GACACAGATTTTCACTCCAGACAAATCCACGGGCGGCTCGAAAAAATTAATGCTGC	1643
OY	1646	GGAATCCAGA---TGTGTGGACACTCCTTATTTATATACAGAGGATCCCGTGTCAAGT	1702
D	1644	GTAACCTGAGTGTAGTGTAGTGTGCTCCTGCTGCTACAGACAAATCCAGAAACTTT	1703
OY	1703	GGGAGTACTGCACCTGACCGCAATGCTCAGACGC	1736
D	1704	ACGACTACTGTGATGTCCCTCAGTGTGGGCCCC	1737
RESULT_14			
LXAX48883			
LOCUS	AY448883	2497 bp	DNA linear PAT 03-JUL-2002
DEFINITION	Sequence 1 from Patent WO0220813.		
ACCESSION	AY448883		
VERSION	AY448883.1	GI:21697781	
KEYWORDS	.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1		
JOURNAL	Materials and methods relating to endothelial cell growth inhibitors		
FEATURES	Patent: WO 0220813-A 1 14-MAR-2002; Karolinska Innovations AB (SE)		
source	Location/Qualifiers		
	1..2497		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	50..2482		
	/note="unnamed product"		
	/codon_start=1		
	/protein_id="CAD37757.1"		
	/db_xref="GI:21697782"		
	/translation="MEHEVVLILLILLSGGEPLDDVYNTOGASLFTVKROLGAG		
	SIRECAKCEDEFEFCRAFOYSKEOCYIAMENRKSIIIRADVLEPKRYISE		
	CKTNGKNYRGTMSKTNGITCOQWSTSPHRPFSPTSGLENYGNNDNDO		
	GPWCYTDPDKRYDYDYLICEBECHMSSENYDGKISKMSGLECOMASOPHAG		
	YTPSKPNKLKKYICNPREDLPKPFETIDPKRMELCIPICTITPPSSGYOOL		
	KGTGENTRAVAVTVSGHTQHSADQPHTHNKIPENPECKNIKDENCYNDGKRAFW		
	CHTNGSVREMYCTIPDCSDSPVSTEDLATPADELPLPVODCYHGGSYRGTSST		
CDS			

BASE COUNT		692	a	624	c	636	g	545	t
ORIGIN									
Query Match	4.7%;	Score 236.6;		DB 6;	Length 2497;				
Best Local Similarity	58.6%;	Pred. No. 3.2e-62;							
Matches 547;	Conservative 0;	Mismatches 315;	Indels 72;	Gaps 5;					
QY	809	AGTGTACCATGATGTAATGACAGAGTAATTGAGGCACATACTCCACACACTGTCAAGGAA	868	TTTGGCCGSSAMPPIRHQKTLPEATPNAGLIMNCRNDAKNGMCFITDVSMEVCA					
DB	870	AGTGTCTGAGGGAAACAGGTGMAAACCTATCGGGGAAATGTGGCTGTGTACCGTCCGGGC	929	LKKSCSTEDASVAPRPYVLLPDEVETPSEEDOMGNGKGYKRAATYTGDSQCDRAAO					
QY	869	GAACCTGGCAAGCTTGGTGATCATCTATGACACACACATGCTCATAGTGGAGCCCAATACT	928	EPHRSILFTPTNPRAGLEENKYNCRNPDGDGEMPCITNPRKLTLDYCDVQCAPSD					
DB	930	ACACCTGTTCAGACTGGAGTGCACAGACCCCTTCACACACATACAGAGGACACCGAATACT	989	CGKQVPEPRKCGRVVGGCVAAHPHSMWVSLRTREFMCGGGTLLISPEWVLLAHL					
QY	929	ACCCAAATGTGGCTTGAATCATGAATCACTACGAGAAATCCAGATGCTGTGGCAGCTCCT	988	EKSPRSPSYVILGAAHVEVLEHPOVETESLRLETFREDLILKLSIPAVLTADKL					
DB	990	TCCCTGCAAAAATTGGATGAAACTACTGCGCCAACTCTGACGAGAAAGAGGCCCAT	1049	PACLPSPNYVADTECFITGMGCTGTFAGGLKEKQLEQVLEIKVKNRSEPLNGRVO					
QY	989	ATTGTTATACGAGGATCCCGGTGTACAGTGGGAATACTGCACCTGACGCAATGCTCAG	1048	STELCAGHAGGTGDSGCGSGSGLVCEKDKYILIGVTSMLGCAAPRNRKGVVYVSR					
DB	1050	GGTGTCCATACAAACCAAGCCAAAGTGGGGATGTAAGTATACGT-----	1100	FVTYIEGVMRN"					
QY	1049	ACGCAAGAGGACCTCCCTCGCGCTCCGAGCTGTATCCCGGGTTTCAAGCCTAGAGGCTC	1108						
DB	1101	-----CCTGTGACTCCTCCCAAGTATCCAG-----GAACAAT	1133						
QY	1109	CTTCCGAAACAGCACCGCACTAGCAAGAGCCTGGSGTCCAGAGTGTCTACATGTAATG	1168						
DB	1134	TGGCTCCCAAGCACCACTGAGCTTAACCCCTGTGTGTCCAGAGACTGTACCATGTGTATG	1193						
QY	1169	GACAGAGTTATGAGGACATATCTCACCACTGTGCAGAGAAAGAACTGCCAAGTTGTGT	1228						
DB	1194	GACAAAGCTACCGAGGACATCTCTCCACACACCACAGAAAGAAAGTGTACGTCTTGGT	1253						
QY	1229	CATCTATGACACCACTGCGATAGTGTGGACCCAGAAATACTACCCAAATGCTGGCTTGA	1288						
DB	1254	CATCTATGACACCAACACGCGGACCCAGAGAACCCCGAAGAACTACCCAAATGCTGGCTTGA	1313						
QY	1289	TCATTAATCTACGCAAGAAATCCAGATGCTGTGGACGCTCCTATATGTAATCGAGAGTTC	1348						
DB	1314	CAATTAATCTACGCAAGAAATCCAGATGCGATTAAGGGCCCTGGGTGTTTACCAAGACCC	1373						
QY	1349	CCGGTGTGAGTGGGAGTACTGTCAACTGTAGCCCAATGCTCAGACGACAGAAAGGATGCGC	1408						
DB	1374	CCAGGGTGAGTGGGAGTACTGTCAACTGTGAAATATGCTCAGAGAACAGAGAGTGTG	1433						
QY	1409	TGCGGCTTCGCACTGTACCCGGGTTCCAAAGCTGTAGAGGCTCTTCCGAAAGAACCGGA	1468						
DB	1434	TAGCACTTCGCGCTTGTGTCTGCTTCCAGATGTAGAAATCTCTTCCGAAAGAA-----	1486						
QY	1469	CTGACCAAAAGGCTGGGTGACAGAGTCTCAATGTAATGAGACAGAGTTATGAGGCA	1528						
DB	1487	-----GACTGTATGTTTGGGAATGGGAAAGATACGAGGCA	1523						
QY	1529	CATTACTCCACCACTGTCAAGAGAACCTGCGCAAGCTTGTGTCTATGACACCACACT	1588						
DB	1524	AGAGGGCGACCACTGTACTGTGGAGCGCATGTAGAGATGTGGCTGCCAGAGACCCCATTA	1593						
QY	1589	CGCATAG---TGGAGCCCAAGAAATCTACCCAAATGCTGGCTGTGATCTGAATCTACTCA	1645						
DB	1584	GACAAAGCATTTTCACTCCAGAGCAAAATCCAGCGGCGGGTCTGGAAAAAAATTTACTGCC	1643						
QY	1646	GGAATTCAGAA--TGTCTGTGGCAGCTCTATTTGTAATGAGAGGATCCGGGTCTCAGGT	1702						

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 00:20:02 ; Search time 888 Seconds
(without alignments)
12680.171 Million cell updates/sec

Title: US-09-923-515-3_COPY_1_5000

Perfect score: 5000
Sequence: 1 cctggattggacacattt.....caagcttgcatcatatgac 5000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq-101002:*

1: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	489.6	9.8	924	22	AAD03255
2	236.6	4.7	1719	20	AAx86804
3	236.6	4.7	1719	20	AAx83345
4	236.6	4.7	2178	23	AAx80436
5	236.6	4.7	2433	20	AAx77711
6	236.6	4.7	2433	24	ABN89459
7	236.6	4.7	2497	21	AAx52284
8	236.6	4.7	2497	24	ABN85296
9	236.6	4.7	2679	14	AA040319

10	236.6	4.7	2732	20	AAx5376	SEQ ID 51 of W0991
11	236.6	4.7	2732	24	ABN81696	Human plasminogen
12	236.6	4.7	2753	12	AAQ12547	Encodes Plasminogen
13	236.6	4.7	2756	12	AAQ12542	Encodes Plasminogen
14	236.6	4.7	2756	12	AAQ12548	Encodes Plasminogen
15	236.6	4.7	2756	12	AAQ12552	Encodes Plasminogen
16	236.6	4.7	2756	12	AAQ12553	Encodes Plasminogen
17	236.6	4.7	2756	12	AAQ12554	Encodes Plasminogen
18	236.6	4.7	2756	23	AAx66112	DNA encoding novel
19	236.6	4.7	2756	23	AAx80439	DNA encoding novel
20	236.6	4.7	2759	12	AAQ12549	Encodes Plasminogen
21	236.6	4.7	2759	12	AAQ12550	Encodes Plasminogen
22	236.6	4.7	2759	12	AAQ12557	Encodes Plasminogen
23	236.6	4.7	2759	12	AAQ12558	Encodes Plasminogen
24	236.6	4.7	2759	12	AAQ12543	Encodes Plasminogen
25	236.6	4.7	2762	12	AAQ12544	Encodes Plasminogen
26	236.6	4.7	2762	12	AAQ12546	Encodes Plasminogen
27	236.6	4.7	2762	12	AAQ12545	Encodes Plasminogen
28	236.6	4.7	2765	12	AAQ12556	Encodes Plasminogen
29	236.6	4.7	2771	12	AAQ12551	Encodes Plasminogen
30	236.6	4.7	2771	12	AAQ12555	Encodes Plasminogen
31	236.6	4.7	6020	11	AAQ06648	Plasminogen gene f
32	235	4.7	2296	14	AAQ40258	Sequence of a DNA
33	235	4.7	2296	11	AAQ40258	Plasminogen DNA an
34	235	4.7	2296	21	AAx89829	Plasminogen DNA an
35	235	4.7	2296	22	AAx12747	Human plasminogen
36	235	4.7	2497	24	ABx54038	Human plasminogen
37	234.6	4.7	6010	12	AAQ11998	Human plasminogen
38	234.6	4.7	2679	18	AAx89686	Plasminogen encodi
39	216.2	4.3	1133	24	ABx8257	Canine pro-angiot
40	216.2	4.3	1134	21	AAx86802	Human angiotatrin
41	216.2	4.3	1154	20	AAx86805	Angiotatrin K1-4E
42	216.2	4.3	1154	20	AAx86806	Angiotatrin K1-4E
43	216.2	4.3	1154	20	AAx35350	Nucleic acid encod
44	206	4.1	2696	20	AAx77712	Murine plasminogen
45	202	4.0	273	22	AAD03256	Human apoliprote

ALIGNMENTS

RESULT 1	
AAD03255	
ID	AAD03255 standard; cDNA: 924 BP.
AC	AAD03255;
XX	
DT	13-JUN-2001 (first entry)
XX	
DE	Human angiogenesis inhibitor, LK68 cDNA.
XX	
KW	Human: angiogenesis inhibitor; LK68; apolipoprotein(a) kringle domain;
KW	angiogenesis-mediated disease; cancer; rheumatoid arthritis; therapy;
KW	cytostatic; antineoplastic; antiarthritic; antipsoriatic; psoriasis;
KW	ocular angiogenic disease; endothelial cell proliferation; tumour;
XX	cell migration; ss.
OS	Homo sapiens.
XX	
PH	key
FT	1..924
FT	location/Qualifiers
FT	CD5
FT	/*tag= a
FT	/product= "Human LK68 protein"
FT	/note= "CD5 does not include start and stop codon"
FT	/partial
PN	W0200119868-A1.
XX	
PD	22-MAR-2001.
XX	
PF	15-SEP-1999;
XX	99WO-KR00554.
PR	15-SEP-1999;
	99WO-KR00554.

XX (MOGA-) MOGAM BIOTECHNOLOGY RES INST.
 PA Chang J, Kim JS, Park EJ, Yun J, Chung S;
 PI WPI: 2001-244787/25.
 XX P-PSDB: AAY72944.
 DR Novel angiogenesis inhibitor, LK68 for treating angiogenesis-mediated
 XX diseases, e.g. cancer and rheumatoid arthritis, has human
 PT apolipoprotein(a) kringle domains IV36, IV37 and V38 amino acid
 PT sequence -

PS Claim 8; Page 41-42; 50pp; English.

CC The present sequence is a cDNA encoding human angiogenesis inhibitor,
 CC LK68 protein. LK68 protein contains the amino acid sequences of human
 CC apolipoprotein(a) kringle domains IV36 (LK6 protein), IV37 (LK7 protein)
 CC and V38 (LK8 protein). LK68, LK6, LK7 and LK8 are inhibitors are of
 CC endothelial cell proliferation, cell migration and normal development of
 CC capillaries in the chick embryo chorioallantoic membrane (CAM). LK68
 CC protein, its single kringles or their functional equivalents, are useful
 CC for treating angiogenesis-mediated diseases, such as cancer, rheumatoid
 CC arthritis, psoriasis or ocular angiogenic disease in animals or humans.
 CC LK68 is useful as an anticancer agent and also for inhibiting primary
 CC tumour growth.

SQ Sequence 924 BP; 256 A; 252 C; 225 G; 191 T; 0 other:

Query Match 9.8%; Score 489.6; DB 22; Length 924;
 Best Local Similarity 83.0%; Pred. No. 3e-142;
 Matches 558; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

OY 107 AAAGCATGTCGTCAGGATTTGCTACCATGTCAGACAGATTTTCAGGACACTACT 166
 DB 2 AAAGCCTGTGTCTCAGGATTTGCTACCATGTCAGACAGATTTTCAGGACACTACT 61
 OY 167 CCACACTGTTCACAGAGAGACCTGCCAGCTGTGTCATGTATGACACCATCAATATA 226
 DB 62 CCACACTGTTCACAGAGAGACCTGTGTCATGTATGACACCATCAATATA 121
 OY 227 ATAGACACAGAAACTACCAATGCTGCTTATGATCAACTAGCAGGAATCCAG 286
 DB 122 AGAGAGACCCAGAAACTACCAATGCTGCTTATGATCAACTAGCAGGAATCCAG 181
 OY 287 ATGCTGTGACAGCTCTTATGTTATATGAGAGAGATCCCGCTGTCAGGTGAGTACTGCA 346
 DB 182 ATTTGTGGAAACACCTGTGTTACACACCATCCGATCGTGTGAGGTGGAGTACTGCA 241
 OY 347 ACCTGACGCAATGCTGACAGCAGAGAGGACTCCGTCGCGCTCCGACTGTATCCCGG 406
 DB 242 ATCTGACACATGCTCAGAAACAGAAATCAAGTGTCTTAGAGACTCCGACTGTGTCCAG 301
 OY 407 TTCCAAAGCTAGAGGCTCTTCGGAACAGCAGCAGTCCGAAAGCCCTGGGGTCTCAG 466
 DB 302 TTCCAAAGCTAGAGGCTCTTCGGAACAGCAGCAGTCCGAAAGCCCTGGGGTCTCAG 361
 OY 467 AGTCTACCATGTTATGACAGAGTTATGAGGACACATCTCCACACTGTCTACAGAA 526
 DB 362 AGTCTACCATGTTATGACAGAGTTATGAGGACACATCTCCACACTGTCTACAGAA 421
 OY 527 GAACCTGCCAAGCTGTGTCATGTCATGACACACATCTCCGATAGTCCGACCCCGAATACT 586
 DB 422 GGACATCTCAATCTGTGTCATGTCATGACACACATCTCCGATAGTCCGACCCCGAATACT 481
 OY 587 ACCCAATGCTGCTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 646
 DB 482 ACCCAATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 541
 OY 647 ATTGTATAGAGGATCCCGGTGTGTCAGGTGGAGTACTGCAACTGACGCAATGCTCAG 706
 DB 542 GGTGTATACAGGACCCGACATGAGGTGAGTACTGCAACTGACGCGATGCTCAG 601

OY 707 ACGCAGAAAGGACCTGCGGCTCCGACTGTACCCCGTTCAGACCTAGAGCTC 766
 DB 602 ACACAGAAAGGACCTGCGGCTCCGACTGTACCTCAGTCTCAAGCTTAGGCGCTC 661
 OY 767 CTTCGCAACAG 778
 DB 662 CTTCTGAAACAG 673

RESULT 2

AXX86804 standard; cDNA; 1719 BP.

AXX86804;
 20-SEP-1999 (first entry)
 Human plasminogen cDNA.

Angiostatin; growth inhibition; endothelial cell; angiogenesis;
 tumour; solubilisation; bacterial inclusion body; plasminogen; ss.

Homo sapiens.

MO9935248-A2.

15-JUL-1999.

11-JAN-1999; 99WO-US00048.

12-JAN-1998; 98US-0071247.

(SEAR) SEARLE & CO G D.

Casperson GF, Polazzi JO, Violand BN;

WPI: 1999-430389/36.

Producing biologically active angiostatin by solubilizing and
 refolding recombinant protein, used as anticancer and
 anti-angiogenic agent

Example 1; Page 47-48; 64pp; English.

The invention describes the production of angiostatin that comprises (a)
 culturing cells expressing a gene encoding angiostatin; (b) recovering
 the gene product; (c) solubilizing and then refolding it at high pH, and
 (d) isolating the properly folded angiostatin. Angiostatin inhibits the
 growth and migration of endothelial cells, and their ability to form
 tubular structures. Angiostatin is used to inhibit growth of tumours
 and angiogenesis. This method provides efficient solubilisation of
 CC angiostatin-containing bacterial inclusion bodies and subsequent
 CC refolding to biologically active protein. Sequences AAX86805-824
 CC represent specific nucleic acid sequences encoding angiostatin that can
 CC be used in the method of the invention. The present sequence represents a
 CC human plasminogen cDNA which acts as a template for PCR reactions to
 CC amplify angiostatin fragments.

SQ Sequence 1719 BP; 507 A; 446 C; 417 G; 348 T; 1 other;

Query Match 4.7%; Score 236.6; DB 20; Length 1719;
 Best Local Similarity 58.6%; Pred. No. 3e-63;
 Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;

OY 809 AGTGTACCATGTTATGACAGAGTTATGAGGACATCTCCACACTGTCTACAGAA 868
 DB 803 AGTGTCTAAGAGAAAGGATGAAATATATGCGGGAATGTGCTGTACGCTGTCCGGCC 892
 OY 869 GAACCTGCCAAGCTGTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 928
 DB 893 ACACCTGTACACTGTGAGTGTGACAGAGCCCTCACAACATTAACAGAGACCCGAAACT 952
 OY 929 ACCCAATGCTGCTGTATGATGAACTAGCAGGAATCCAGATGCTGTGCGAGCTCTT 988

Db 1217 CATCTATGACACCAACCGGACACCAAGACCCAGAAACCTCAATCTGCGCTGA 1276
QY 1289 TCATGAATCTCTCAGAGATCCAGATGCTGTGGACCTCTTATTGTTATACGAGGATC 1348
Db 1277 CAATGAATCTCTCAGAGATCCAGATGCTGTGGACCCCTGCTTTTACACAGACC 1336
QY 1349 CCGGTGTGAGTGGAGTACTGACATGACGCAATGCTCAGACGACAGAGGACCTGCG 1408
Db 1337 CCAGCGTCAAGTGGAGTACTGCAATGCAAAAATGCTCAGAGAACAGACGAGTGTG 1396
QY 1409 TCGGCGCTCGAGCTGTACCCGGTTCCAGGCTTACAGGCTCTCTTCGACACAGACCGA 1468
Db 1397 TAGACCTCGCGCTGTGCTGCTTCAGATAGACACACCTCTCCAGAA----- 1449
QY 1469 CTGAGCAAAAGCGCTGGGAGTGCAGAGTGTACATGATAGACAGAGATTATGAGCA 1528
Db 1450 -----GACTGATGTTTGGGAATGGGAAAGATACCGAGCA 1486
QY 1529 CATATCTCACACTGTACAGAAAGACCTGCCAAGCTTGTCTATGTACACCACT 1588
Db 1487 AGAGGGGACCACTGTACTGGAGCCCATGCCAGGACTGGGCTGCCAGGAGCCCATATA 1546
QY 1589 CGCATAG---TCGACCCCGAGATACTACCAATGCTGGCTGTATCATGACTACTGCA 1645
Db 1547 GACACAGCATTTTCACTCCAGAGACAAATCCAGGGCGGCTGTGAAAAAATTACTGCC 1606
QY 1646 GGAATCCAGA---TCGTGTGCACTCTTATTGTTATACGAGGATCCGGTGTACAGT 1702
Db 1607 GTAACTGTATGTATGATGAGTGTGCTGCTGTACAGCAATCCAAAGAAACTTT 1666
QY 1703 GGGAGTACTGCACTGACGCAATGCTGCAGAGC 1736
Db 1667 ACGACTACTGTATGTCTCTCATGTGCGGCCCC 1700

RESULT 4
AAS80436/c
ID AAS80436 standard; cDNA: 2178 BP.
XX
AC AAS80436;
XX
XX 13-FEB-2002 (first entry)
DT
XX
XX DNA encoding novel human diagnostic protein #16240.
DE
XX
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS
XX
XX Homo sapiens.
PN WO200175067-A2.
XX
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001; 2001WO-US08631.
PF
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Drmanac RT, Liu C, Tang YT;
PI
XX
XX WPI: 2001-639362/73.
DR P-P-SDB; ABG16249.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1: SEQ ID No 16240; 103bp; English.

XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers) and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2178 BP; 453 A; 532 C; 574 G; 619 T; 0 other;

Query Match 4.7%; Score 236.6; DB 23; Length 2178;
Best Local Similarity 58.6%; Pred. No. 3.4e-63;
Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;

QY 809 AGTGTACCATGATGATAGACAGAGATATGAGCGCATATCTCCACATGTCACAGGAA 868
Db 1250 AGTGTCTAAGGGAACAGGGAATATATCGGGATGATGGGTGTATCCGTGTCGGGC 1191
QY 869 GAACCTGCGAAGCTTGTGCTATATGACACCACTCGCATAGTGGACCCAGAAATCT 928
Db 1190 ACACCTGTACAGCTACTGATGATGATGATGATGATGATGATGATGATGATGATG 1131
QY 929 ACCCAATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 988
Db 1130 TCCCTGCAAAAATTTGAGTAAATGATGATGATGATGATGATGATGATGATGATG 1071
QY 989 ATTGTTATACAGAGGATCCCGGTGTGAGTGGAGTATGCAACCTGACGCAATCTCAG 1048
Db 1070 GGTGCTATCAACCAACAGCCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1020
QY 1049 ACGCAGAGGAGTCCGCTGCTCCGCTTCACTGATGATGATGATGATGATGATGATG 1108
Db 1019 -----CTGTGATCTCTCCAGATATGACG-----GAACAAT 987
QY 1109 CTTCGGAACAGCAGCAGTACGAAAGCGTGGGTGAGAGAGTACATGATGATG 1168
Db 986 TGGCTCCACAGACACACCTGAGCTAACCCCTGTGTCCAGAGACTGCTACATGATG 927
QY 1169 GACAGAGTTATGAGCAGCATACTCCACACTGTCACAGAGAAACCTGCCAAGCTTGGT 1228
Db 926 GACAGAGTACGAGGACATCTCCACACACACAGAGAAAGAGTGTGATGTTGGT 867
QY 1229 CATCTATACACCACTCCGATAGTGGACCCCAAAATCTACCAATGTTGGCTTGA 1288
Db 866 CATCTATACACCACTCCGAGCAGAGAGCCCAAAATCTACCAATGTTGGCTTGA 807
QY 1289 TCATGAATCTCTCAGAGATCCAGATGCTGTGGACCTCTTATTGTTATACGAGGATC 1348
Db 806 CAATGAATCTCTCAGAGATCCAGATGCTGTGGACCCCTGCTTTTACACAGACC 747
QY 1349 CCGGTGTGAGTGGAGTACTGCAACCTGACGCAATGCTCAGACGACAGAGGACTGCGG 1408
Db 746 CCAGGTGAGTGGAGTACTGCAACCTGAAAAAAGCTCAGAGACAGAGGAGATGTTG 687
QY 1409 TCGGCGCTCGAGCTGTACCCGGTTCCAGGCTTACAGGCTCTCTTCGACACAGACCGA 1468
Db 686 TAGCACTCGCGCTGTGCTCTTCCATGATGATGATGATGATGATGATGATGATGATG 634

QY 1469 CTGAGCAAAAGCCCTGGGGTGCAGAGTCTACATGTTAATGACAGAGTTATCGAGCA 1528
Db 633 -----GACTGTATGTTGGGAATGGGAAAGATACCGAGGCA 597
QY 1529 CATACTCCACACCTGTGACAGAAAGAACCTGCCAAGTTGTGATCTATACACACACT 1588
Db 596 AGAGGGGACACCTGTTACGTGGAGCGCATGCCAGGACTGGCTGCCAGAGACCCCATATA 537
QY 1589 CGCATAG---TCGGACCCCAAGATACCAATGCTGGCTTGATCAGTACTGCA 1645
Db 536 GACACAGCATTTTACCTTCACAGACAAATCCACGGCGGGCTGAGAAAAAATTACTGCC 477
QY 1646 GGATCCAGA---TGTGTGGACACTCTTATTTATAGAGGATCCCGGTGTCAAGT 1702
Db 476 GTACCCCTGATGTGATGTAGTGTGCTCCGTGTGTACAGACAAATCCAAAGAACTTT 417
QY 1703 GGGAGTACTGCACCTGCAGCAATGCTCAGACG 1736
Db 416 AGGACTACTGTGATGTCTCCTCAGTGTGGGCC 383

RESULT 5
AA77711
ID AA77711 standard; DNA; 2433 BP.
XX
AC AA77711:
XX 10-AUG-1999 (first entry)
DE Human plasminogen DNA coding region.
XX
KW Plasminogen; human; angiotatin; endostatin; gene therapy; vector;
KM anti-angiogenic; attenuation; cytosstatic; anti-diabetic; ophthalmology;
KM tumour growth; solid tumour; diabetic retinopathy; retina; ss.
XX
OS Homo sapiens.
XX
PN W0926480-A1.
XX
PD 03-JUN-1999.
XX
PF 20-NOV-1998; 98WO-US24950.
XX
PR 20-NOV-1997; 97US-0975424.
XX
PA (GENE-) GENETIX PHARM INC.
PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX
PI Bachelet T, Leboulch P, Pawlilk RJ;
XX
DR WPI: 1999-357696/30.
DR P-PSDB; AAY08685.
XX
PT Anti-angiogenic gene therapy vectors
XX
PS Disclosure; Fig 5; 83pp; English.
CC This invention describes a novel viral gene therapy vector comprising a
CC nucleic acid molecule encoding an anti-angiogenic polypeptide chosen
CC from human or murine angiotatin, human or murine endostatin and
CC angiogenesis-inhibiting fusions and fragments, where the viral vector is
CC sufficiently attenuated for use in human gene therapy. The products of
CC the invention have anti-angiogenic, cytosstatic, anti-diabetic and
CC ophthalmological activity. The vector is used in gene therapy for
CC inhibiting tumour growth in humans harbouring a solid tumour. The vector
CC expresses an anti-angiogenic polypeptide. An additional use comprises
CC treatment of diabetic retinopathy, where the anti-angiogenic polypeptide
CC inhibits angiogenesis in the vicinity of the retina. The vector is
CC administered to cells ex vivo and then administered to the patient.
XX
SQ Sequence 2433 BP; 678 A; 607 C; 616 G; 532 T; 0 other;

Query Match 4.7%; Score 236.6; DB 20; Length 2433;

Best Local similarity 58.6%; Pred. No. 3,5e-63;
Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;
QY 809 AGTGTACCATGTATGATGACAGAGTTATCGAGGACATATCCACCATGTCACAGAA 868
Db 821 AGTGTCTGAGGAGAACAGAGTAAACTATCGCGGAATGTGGCTTTACCGTTCCGGC 880
QY 869 GAACCTGCCAAGCTTGTCTATCTATGACACACACTGCGATATGTGGAACCCAGATACT 928
Db 881 ACACCTGTACAGCATGTGAGTGACAGACCCCTCACACATACAGACACACAGAAACT 940
QY 929 ACCCAATGTGGCTGTATGATGAACTACTGCAGGAATCCAGATGTGTGGACCTCTT 988
Db 941 TCCCTGCAAAATTTTGATGAAACTACTGCGCATCTGACAGGAAAAAGGCCCAT 1000
QY 989 ATTGTATACGAGGATCCCGGTGTGAGTGGAGTACTGCACCTGACGATGTCTAG 1048
Db 1001 GGTGCCATFCAACAACAGCCAAAGTGGGTGGAGTACTGTAAATATACCTT----- 1051
QY 1049 AGCAGAAAGGACCTGCGGCTCGAGCTTACCGGTTTCCAAAGCCTAGAGGCTC 1108
Db 1052 -----CCTGTACTCTCCCATATACAG-----GAACAAT 1084
QY 1109 CTTCGACACACACACGACTAGACAAAGCCTGGGTGTCAGAGTGTATCCATGTAATG 1168
Db 1085 TGGCTCCACACACACACACTGTAGTAACCTGTGTGTCAGAGTGTATCCATGTGATG 1144
QY 1169 GACAGATTATGAGGACATATCTCCACACTGTACACAGAAAGAACCTGCCAAGTGTGT 1228
Db 1145 GACAGGCTACCGAGGACATCTCCACACACACACAGGAAAGAGTGTAGCTGTGT 1204
QY 1229 CATCTATGACACACACACATGCTATGTCGACCCCGAATATCTACCAATGCTGCTTGA 1288
Db 1205 CATCTATGACACACACACGCGACACAGAAAGCCCGAATATCTACCAATGCTGCTGA 1264
QY 1289 TCATGAACTACTGACAGGAATCCAGATGCTGTGGACGCTCTTATTTATACAGGATC 1348
Db 1265 CATGAACTACTGACAGGAATCCAGATGCGATTAAGGCCCCCTGTGTTTACCAACAC 1324
QY 1349 CCGGTGTGAGGGGAGTACTGCAACCTGACGCAATGCTCAGACAGCAAGAGACTCGG 1408
Db 1325 CCAGGCTCAGGGGAGTACTGCAACCTGAAATGCTCAGAGACAGACAGAGTGTG 1384
QY 1409 TCGCGCTCCGACTGTACCCCGGTTCAGAGCTTAGAGCTCTTCCGACACAGACCGA 1468
Db 1385 TAGCACCCTCGGCTGTGTCTGTCTCCAGATGTAGAGATCTTCCGAGAA----- 1437
QY 1469 CTGAGCAAAAGCCCTGGGGTGCAGAGTCTACATGTTAATGACAGAGTTATCGAGCA 1528
Db 1438 -----GACTGTATGTTGGGAATGGGAAAGATACCGAGGCA 1474
QY 1529 CATACTCCACACCTGTGACAGAAAGAACCTGCCAAGTTGTATCTATGACACACT 1588
Db 1475 AGAGGGGACACACTGTATGAGGAGCGCATGCGAGACTGGGCTGCCAGAGGCCCATATA 1534
QY 1589 CGCATAG---TCGGACCCCAAGATACCAATGCTGGCTTGATCATGTAACACTGCA 1645
Db 1535 GACACAGCATTTTACCTTCACAGACAAATCCAGCGCGGGCTGAGAAAAAATTACTGCCC 1594
QY 1646 GGATCCAGA---TGTGTGGACAGCTCTTATTTATAGAGGATCCCGGTGTCAAGT 1702
Db 1595 GTACCCCTGATGTGATGTAGTGTGCTCCGTGTGTACACAGCAATCCAAAGAACTTT 1654
QY 1703 GGGAGTACTGCACCTGCAGCAATGCTCAGACG 1736
Db 1655 ACGACTACTGTGATGTCTCCTCAGTGTGGGCC 1688

RESULT 6
ABN89459
ID ABN89459 standard; cDNA; 2433 BP.
XX
AC ABN89459;


```

FT      /*tag= a
FT      /product= "Human plasminogen"
FT      /transl_except= (pos:2150..2152, aa:Ile)
XX
XX      US6057122-A.
XX
XX      02-MAY-2000.
XX
XX      05-MAY-1997; 97US-0851350.
XX
XX      03-MAY-1996; 96US-0643219.
XX      03-APR-1997; 97US-0832087.
XX
XX      (ABBO ) ABBOTT LAB.
XX
XX      Davidson DJ;
XX
XX      WPI: 2000-349573/30.
XX      P-PsDB; AAB01887.
XX
XX      Preparation of Kringle five peptide fragment for treating various
XX      disorders such as angiogenic, ocular, skin diseases and cancer,
XX      involves mixing mammalian plasminogen and elastase followed by
XX      incubation and isolation -
XX
XX      Disclosure: Fig 3; 48pp; English.
XX
XX      The invention relates to a method of preparing plasminogen kringle 5
XX      peptide fragments. The method comprises mixing mammalian plasminogen and
XX      elastase in the ratio 1:100-1:300, followed by incubating and isolating
XX      the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
XX      endothelial cell proliferation and migration. The peptides are useful for
XX      treating angiogenic diseases, primary and metastatic solid tumours and
XX      carcinomas of various organs such as breast, genital tract, endocrine
XX      glands, skin, tumours of the brain and eyes and solid tumours arising
XX      from haematopoietic malignancies such as leukemias and lymphomas. They
XX      are also used for the prophylaxis of various autoimmune diseases (e.g.,
XX      rheumatoid arthritis), ocular diseases, skin diseases (e.g., psoriasis),
XX      blood vessel diseases (e.g., haemangiomas, Osler-Weber Syndrome),
XX      diseases caused by excessive or abnormal stimulation of endothelial cells
XX      (e.g., Crohn's disease, atherosclerosis), diseases which have
XX      angiogenesis as a pathologic consequence (e.g., cat scratch disease and
XX      ulcers). The peptides are also useful as a birth control agent which
XX      inhibits ovulation and establishment of the placenta. The present
XX      sequence represents DNA encoding human plasminogen.
XX
XX      Sequence 2497 BP; 691 A; 624 C; 637 G; 545 T; 0 other;
XX
XX      Query Match 4.78; Score 236.6; DB 21; Length 2497;
XX      Best Local Similarity 58.68; Pred. No. 3; 6e-63;
XX      Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;
XX
XX      809 AGTGTACCATGTAATGACAGAGTATGAGGACATATCTCCACCTGTCCAGGAA 868
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      870 ACTGTCTGAAGGAACAGTGAACAACTATCGGCGAATGCGTGTATACCTGTCCGGGC 929
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      866 GAACCTGCCAAGCTTGGTCACTATGACACCACTCGCTAGTGGAGCCCAAGTACT 928
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      930 ACACCTGTACGACTGCGAGTGACAGACCCCTCACACATTAACAGAACCAAGAAACT 989
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      929 ACCCAAAATGTGCTGTATGATCACTGACAGAGATCCAGATGCTGGGAGCTCCTT 988
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      990 TCCCTGCAAAAATTTGGATGAATAAATCTATGCGCAATCTCTGACGAAAAAGGCCCAT 1049
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      989 ATTGTTATACGAGGAGATCCCGGTCTCAGGTGGAGTACTGCAACCTGAGCGCAATGCTCAG 1048
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1050 GGTGGCATATACCAACCAAGCCAAAGTGGGAGTACTGTAGATATACGCT----- 1100
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1049 ACGGAGAGGAGTCCCTCGCGCTGCACTGTATACCGCGGTTCCAGGCTTACAGAGCTC 1108
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1101 -----CCTGTGACTCTCTCCCGAGTATCCAG-----GAACAT 1133
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1109 CTTCCGAACAGCACCACTGACCAAGGCTGGGGTGACGAGGTCTTACATGTATATG 1168

```

```

DB      1134 TGGCTCCACAGACCACTGAGCTAACCCTGTGCTCCAGAGCTGTACCATGTGATG 1193
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1169 GACAGAGTATGAGGACACATATCTCAGCAGTGTACAGGAAAGAACTGCCAAGTTGCT 1228
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1194 GACAGAGCTACGAGGACATCTCCACACACACAGGAAAGTGTAGCTTGGT 1253
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1229 CATCTATGACACACACTGCTAGTCCGACCCCAAGATATACCAAAATGTGGCTTGA 1288
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1254 CATCTATGACACACACCGGACACAGAACCCCGAAGAAATATACCAAAATGTGGCTGA 1313
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1289 TCAATGACTACTGAGGAATCCAGATGCTGTGGAGCTCTTATTTATACAGGATC 1348
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1314 CAATGAACTACTGACAGAAATCCAGATGCGGATTAAGGATTAACACAGACCC 1373
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1349 CCGGTGTACAGTGGAGTACTGCACTGACAGCAATGCTCAGACGAAAGGAGCTCCG 1408
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1374 CCAAGCTCAGGTGGAGTACTGCAACCTGMAAAATGTCTAGAACAGAACGAGTCTTG 1433
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1409 TCGCGCTCCGACTGTATACCCCGGTTCCAAGCTTAGAGGCTCTCCGAACAGACCGA 1468
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1434 TAGACCTCCGCTGTTGTCTGCTCCAGATGTAGAGACTCTTCGGAAGAA----- 1486
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1469 CTGAGCAAGGCGCTGGGGTGCAAGAGTGTACCATGTATAGACAGATTTATGAGGCA 1528
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1487 -----GACTGTATGTTGGGAATGGAAAGATCCGAGGCA 1523
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1529 CATATCCACACTGTACAGAGAAAGACCTGCCAAGCTTGTCTATATACACCACT 1588
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1524 AGAGGGGACACACTGTATCTGTGGAGCGCATGTCAGAGTGGGCTGCCAGAGGCCATA 1583
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1589 CGATATG---TCGGACCCCAAGATATACCCAAATGCTGGCTGTATCATGAACTAGCA 1645
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1584 GACACGACATTTTCACTCCAGAGAACATCCAGGCGGCTGTGAAAAAATTTACTGCC 1643
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1646 GGAATCCAGA---TGCTGTGCAAGCTTATTTATTTATAGAGGATCCGGTGCAGT 1702
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1644 GTAAACCTGTATGTATGTAGTGTGGTCCGTGTCTACACAGCAAAATCCAAAGAACTTT 1703
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1703 GGGAGTACGTCAACCTGACGCAATGCTCAGACGC 1736
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      1704 ACGACTACTGTATGTCTCTCATGTGCGGCC 1737
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
ABN85296
ID ABN85296 standard; DNA; 2497 BP.
XX
XX      AC ABN85296;
XX
XX      30-SEP-2002 (first entry)
XX
XX      Human plasminogen coding sequence.
XX
XX      DE Human; antirheumatic; antiarthritic; gene therapy; anti-angiogenic;
XX      KW rheumatoid arthritis; plasminogen; angiotatin; gene; ds.
XX
XX      OS Homo sapiens.
XX
XX      FH Key Location/Qualifiers
XX      FT CDS 50..2482
XX      FT /tag= a
XX      FT /product= "Human plasminogen"
XX      FT 50..103
XX      FT /tag= b
XX
XX      PD 11-JUL-2002.
XX      PD WO200253191-A1.
XX
XX      PE 03-JAN-2002; 2002WO-KR00001.
XX      PE 05-JAN-2001; 2001KR-0000691.

```



```
QY 1049 ACGCAGAAAGGACTGCGCTGCGCCTCCGACTGTTACCCCGGTTCCAGACCTAGAGGCTC 1108
Db 1106 -----CCGTGACTCTCTCCCAAGTATCCAGC-----GACAAAT 1138
QY 1109 CTTCCGACACAGACGACGACTGAGCAAAAGGCTGGGGGTGCAGAGTGTCTACCATGGTATG 1168
Db 1139 TGGCTCCACAGCAGCAGCTGAGCTTAACCCCTGTGCTCCAGACTGTACCATGGTATG 1198
QY 1169 GACAGATTTATCGAGCAGATATCTCCACCAGCTGTACAGAGAAAGAACTGCCAAGCTTGGT 1228
Db 1199 GACAGAGCTTACCGAGGACATCTCTCCACCAGCAGAGAAAGAAAGTCTGAGTCTTGGT 1258
QY 1229 CATCTATGACACACACTGCGATAGTGGACCCAGAAATACTACCCAAATGCTGGCTTGA 1268
Db 1259 CATCTATGACACACACCCGACACAGAAAGACCCAGAAATCTGACCAATCTGGCCCTGA 1318
QY 1289 TCATGAACTACTGACGAAATCCAGATGCTGTGCGACGCTCTTATTTATATGAGAGATC 1348
Db 1319 CAATGAACCTACTGACGAAATCCAGATGCGGATTAAGGCCCTGTGTTTATCCACAGAC 1378
QY 1349 CCGGTGTACAGTGGAGTACTGTACACCTGACCAATGTCTCAGAGCGCAGAAAGGACTGCCG 1408
Db 1379 CCAGCTGACAGTGGAGTACTGTCAACCTGAAAAAATGCTCAGAGAACAGAAAGGAGTGTG 1438
QY 1409 TCGGCGCTCGGACTGTTAACCCCGGTTCCAAAGCCTAGAGGCTCTCTTATTTATGAGAG 1468
Db 1439 TAGACCTCTCGGCTGTGTTCTCTCTCCAGATGTAGAGACTCTTCTCCAGAA----- 1491
QY 1469 CTGAGCAAAAGGCTGGGGGTGCAGAGTGTCTACCATGTTAATGAGACAGAGTTTCAGAGCA 1528
Db 1492 -----GACTGTATGTTTGGGAATGGGAAAGATACCGAGCA 1528
QY 1529 CATACTCCACACTGTCACAGAAAGAACTGGCAAGCTTGCATCTATGACACACACT 1588
Db 1539 AGAGGGGACCACTGTACTAGGAGACGCCATGCGAGACTGGGCTGCCAGAGGCCCACTA 1588
QY 1589 CGCATAG---TCGACCCCGAATACTACCCAAATGCTGGCTTGATCATGAACTACTGCA 1645
Db 1589 GACACAGCATTTTCACTCCAGAGAACAAATCCACGGCGGTCTGAAAAAATTTACTGCC 1648
QY 1646 GGAATCCAGA---TGCTGTGGAGCTCTTATTTATGAGAGGATCCCGGTGTGCAAGT 1702
Db 1649 GTAAACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1708
QY 1703 GGGAGTACTGCAACCTGACGCAATGCTCAGAGCG 1736
Db 1709 ACGACTACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1742

RESULT 12
AAQ12547
ID AAQ12547 standard; DNA; 2753 BP.
XX
XX AAQ12547;
AC
AC
DT 23-SEP-1991 (first entry)
XX
XX Encodes Plasminogen mutuin T1 with thrombin cleavage site.
XX
XX protease; fibrinolysis; blood clotting; ss.
XX
XX Key Location/Qualifiers
XX mutation 1796..1801
XX /*tag= a
XX /*tag= "replaces CCTGGA (Pro-Gly) with GGTCTT
XX /*tag= b
XX /*tag= modified plasminogen
XX
XX W09109118-A.
XX
XX PD 27-JUN-1991.
```

```
XX
PF 07-DEC-1990; 90WO-G001912.
XX
PR 07-DEC-1989; 89GB-0027722.
XX
PR 07-DEC-1990; 90WO-G801911.
XX
PA (BRB1-) BRIT Bio-TECHN LTD.
XX
XX Dawson KM, Edwards RM, Forman JM;
XX
XX WPI: 1991-208145/28.
XX
XX P-PSDB: AAR12938.
XX
XX Activatable fibrinolytic and antithrombotic proteins - activated by
XX e.g. factor Xa, thrombin or activated protein C
XX
XX Claim 16; Fig 2 and Fig 5; 73pp; English.
XX
CC This sequence encodes a plasminogen mutant cleavable by thrombin.
CC Activation is localised to the thrombus because cleavage to plasmin
CC is by an enzyme of the blood clotting pathway. Compositions
CC comprising the mutant plasminogen encoded by this sequence are used
CC for treatment or prevention of thrombosis, etc.
CC See AAQ12542-Q12558.
XX
XX
SQ Sequence 2753 BP; 764 A; 671 C; 695 G; 623 T; 0 other;

Query Match 4.7%; Score 236.6; DB 12; Length 2753;
Best Local Similarity 65.6%; Pred. No. 3,7e-63;
Matches 410; Conservative 0; Mismatches 179; Indels 36; Gaps 3;

QY 92 CAGCAGACCTGAGCAAGGCCATGTGTCTCAGATGTGCTACCATGTGATGACAGAGTT 151
Db 1158 CAGCAGACCTGAGCAAGGCCATGTGTGTGTCAGAGCTGCTACCATGTGATGACAGAGCT 1217
QY 152 ATGAGAGGACACGATCTCCACCACTGTCCAGAGAAAGACCTGGCAAGTGTGATATGA 211
Db 1218 ACCGAGGACATCTCTCCACCACTGTCCAGAGAAAGAGTGTGATGTGATATGA 1277
QY 212 CACCACTACATATATATAGACACAGAAACTACCAAAATGCTGGCTTATCATGAACT 271
Db 1278 CACCACTACGAGACACAGAAAGACCCAGAAACTACCAAAATGCTGGCTTATCATGAACT 1337
QY 272 ACTGCAAGAAATCCAGATGCTGTGCAAGCTCTTATTTATGACAGAGATCCGGGTGCA 331
Db 1338 ACTGCAAGAAATCCAGATGCTGTGCAAGCTCTTATTTATGACAGAGATCCGGGTGCA 1397
QY 332 GGTGGAGTACTGCAACCTGACGCAATGCTCAGAGCAAGAGGACTGCCGCTC 391
Db 1398 GGTGGAGTACTGCAACCTGACGCAATGCTCAGAGCAAGAGGACTGCCGCTC 1457
QY 392 CGACTGTTAACCCTGGTTCCAAAGCTCTCTTCCGAACAGCAACCGACTGAGCAAA 451
Db 1458 CGCTGTGTCCTGCTTCCAGATGTAGAGACTCTTCCGAAGAA----- 1501
QY 452 GGCCTGGGGTGCAGAGGCTCTACATGGTATGAGACAGATTTACGAGGACATACTCA 511
Db 1502 -----GACTATATGTTTGGAAATGGAAAGATACCGAGCAAGAGGCGCA 1547
QY 512 CCACTGTACAGGAGAAAGCACTGCAAGCTTGGTCTATATGACACACATCGCATAG-- 559
Db 1548 CCACTGTATTTGGAGCGCATGCGAGAGATGGGCTGCCAGAGACCCATAGACACAGCA 1607
QY 570 -TGGAGCCCAAGATATCTACCAATGCTGGCTGTATATGATACATGACGAGAGATCAG 628
Db 1608 TTTTCACTCCAGAGCAAAATCCAGGGGGGCTGTGAAAAAATTTATGCGGTAACTTG 1667
QY 629 A---TGCTGTGAGAGCTCTTATTTATGAGAGATCCCGGTGTCTAGGTGGAGTACT 685
Db 1668 ATGTGATGTAGTGTGTCTCTGTCTACAGCAAAATCCAAAGAAACTTTACGACTACT 1727
QY 686 GCAACTGACGCAATGCTCAGAGCG 710
```

Db 1728 GTGATGTCCTCAGTGTGCGGCCCC 1752

RESULT 13

ID AAQ12542 standard; DNA; 2756 BP.

AC AAQ12542;

DT 23-SEP-1991 (first entry)

DE Encodes Plasminogen mutelin XI with factor Xa cleavage site.

KW protease; fibrinolysis; blood clotting; ss.

FT Key Location/Qualifiers

FT mutation 1796..1801

FT CDS /tag= a /note= "replaces CCT (Pro codon) with two codons"

FT /tag= b /product= modified plasminogen

PN MO9109118-A.

PD 27-JUN-1991.

PF 07-DEC-1990; 90WO-GD01912.

PR 07-DEC-1989; 89GB-0027722.

PR 07-DEC-1990; 90WO-GB01911.

PA (BRBI-) BRIT BIO-TECHN LTD.

PI Dawson KM, Edwards RM, Forman JM;

DR WPI; 1991-208145/28.

DR P-PSDB; AAR12933.

PT Activatable fibrinolytic and antithrombic proteins - activated by

PT e.g. factor Xa, thrombin or activated protein C

PS Claim 16; Fig 2 and Fig 4; 73pp; English.

CC This sequence encodes a plasminogen mutant cleavable by Factor Xa.

CC Activation is localised to the thrombus because cleavage to plasmin

CC is by an enzyme of the blood clotting pathway. Compositions

CC comprising the mutant plasminogen encoded by this sequence are used

CC for treatment or prevention of thrombosis, etc.

CC See AAQ12543-Q12558.

CC

CC

CC

CC

CC

CC

CC

CC

CC

QY 332 GGTGGAGTACTGCAACCTGACGCAATGCTCAGACGAGAAAGGACTGCCGTGGGCGCTC 391

Db 1398 GGTGGAGTACTGCAACCTGCAAAAATGCTCAGAAACAGAGGACTGTTGACACTTC 1457

QY 392 CGACTGTTACCCGCTTCCAAAGCCCTAGAGGCTCTTCCGACAAAGCCGACTGAGCAA 451

Db 1458 CGCTGTGTCCTGCTCCAGATGTAGAGACTCTTCCGAGAA----- 1501

QY 452 GGCCTGGGCTGACGAGTGCATCATGTTATGAGACAGATTATGAGGACATACCTCA 511

Db 1502 -----GACTGATGTTTGGATGGAATGGAAAGGATACGAGGCAAGAGCGGA 1547

QY 512 CCAGTGTCAAGAGAAACCTGCCAAGCTTGCTATCTATGACACACACTGCGATAG-- 569

Db 1548 CCAGTGTACTGGAGCGCATGCGAGGCTGGCTCCGACGAGGCGCCCTATGACACACA 1607

QY 570 -TCGAGCCCGCAATATACCAAAATGCTGATGATGAACTCTCAGAAATCCAG 628

Db 1608 TTTTCAGTCCAGAGCAAAATCAGGCGGCTCTGGAATAATTTAGCTGTAACCTTG 1667

QY 629 A---TCTGTGGCAGCTCTTATTGTTATACGAGGATCCCGGTGTCAGGTGGAGTACT 685

Db 1668 ATGATGATGTAGTGTGCTGCTGCTGCTACGACAAATCCAAAGAAATCTTACGACTACT 1727

QY 686 GCAACCTGACCGCAATGCTCAGAACCC 710

Db 1728 GTGATGTCCTCAGTGTGCGGCCCC 1752

RESULT 14

ID AAQ12548 standard; DNA; 2756 BP.

AC AAQ12548;

DT 23-SEP-1991 (first entry)

DE Encodes plasminogen mutelin T2 with thrombin cleavage site.

KW protease; fibrinolysis; blood clotting; ss.

FT Key Location/Qualifiers

FT mutation 1796..1801

FT /tag= a /note= "replaces CCTGGA (Pro-gly) with GGTGCTCA

FT /tag= b /note= (gly-gly-Pro)"

FT CDS 65..2500

FT /tag= b /product= modified plasminogen

PN MO9109118-A.

PD 27-JUN-1991.

PF 07-DEC-1990; 90WO-GD01912.

PR 07-DEC-1989; 89GB-0027722.

PR 07-DEC-1990; 90WO-GB01911.

PA (BRBI-) BRIT BIO-TECHN LTD.

PI Dawson KM, Edwards RM, Forman JM;

DR WPI; 1991-208145/28.

DR P-PSDB; AAR12939.

PT Activatable fibrinolytic and antithrombic proteins - activated by

PT e.g. factor Xa, thrombin or activated protein C

PS Claim 16; Fig 2 and Fig 5; 73pp; English.

CC This sequence encodes a plasminogen mutant cleavable by thrombin.

CC Activation is localised to the thrombus because cleavage to plasmin

CC is by an enzyme of the blood clotting pathway. Compositions comprising the mutant plasminogen encoded by this sequence are used for treatment or prevention of thrombosis, etc.

CC See AA012542-Q12558.

XX Sequence 2756 BP; 765 A; 671 C; 697 G; 623 T; 0 other;

Query Match 4.7%; Score 236.6; DB 12; Length 2756;
Best Local Similarity 65.6%; Pred. No. 3.7e-63;
Matches 410; Conservative 0; Mismatches 179; Indels 36; Gaps 3;

```
OY 92 CAGCAGCAGCTGACGAAAGCAGTGGTCCAGATGCTACCAATGGTGTGACAGAGTT 151
DB 1138 CAGCAGCAGCTGACGAAAGCAGTGGTCCAGATGCTACCAATGGTGTGACAGAGTT 1217
OY 152 ATCGAGCAGCTGCTCCAGCAGCTGTCCAGAGAGAGCTCCCAAGCTGTGATCATGA 211
DB 1218 ACCGAGGACATCTCCACACCCACACAGGAAAGAGTGTCACTGTGATCATGA 1277
OY 212 CACCATCAACATATATAGGACACAGAAATACCAATGGTGTGATCATGA 271
DB 1278 CACCATCAACATATATAGGACACAGAAATACCAATGGTGTGATCATGA 1337
OY 272 ACTGAGGAAATCAGATGCTGTGAGCAGCTCTTATTTATACGAGGAGTCCGGGTGCA 331
DB 1338 ACTGAGGAAATCAGATGCTGTGAGCAGCTCTTATTTATACGAGGAGTCCGGGTGCA 1397
OY 332 GGTGGAGTACTGCAACCTGACGCAATGCTCAGACAGCAGAGGAGTCCGCGGCTC 391
DB 1398 GGTGGAGTACTGCAACCTGACGCAATGCTCAGACAGCAGAGGAGTCCGCGGCTC 1457
OY 392 CGACGTGATACCCGCTTCCAGAGCTGCTCCGAAACAGCAGCAGTGAAGCA 451
DB 1458 CGCCTGTGCTCTGCTCCAGATGTAGAGCTCTCCGAAAGCA 1501
OY 452 GGCCTGGGTGACGAGTGTCTACCATGTATATGACAGAGTATTCAGAGCAGATCTCA 511
DB 1502 -----GACTGTATGTTTGGAGATGGGAAAGATCCAGCAGAGAGGCGCA 1547
OY 512 CCAGTGTACAGGAAAGCCTGCCAAGCTTGATCATATGACACAGCAGTGCATAG-- 569
DB 1548 CCAGTGTACAGGAAAGCCTGCCAAGCTTGATCATATGACACAGCAGTGCATAG-- 1607
OY 570 -TCGAGCCCAAGATATCTACCAATGCTGTGATCATGATGATCTGAGAGATCCAG 628
DB 1608 TTTTACCTCCAGAGCAATTCAGGAGGCTGTGAAAAAAATTTACTGCGTAACTCTG 1667
OY 629 A---TGCTGTGAGCTCCTTATTTATATGAGAGGATCCCGGTGTCAGTGGAGTACT 685
DB 1668 ATGCTGTGAGCTCCTTATTTATATGAGAGGATCCCGGTGTCAGTGGAGTACT 1727
OY 686 GCAACCTGAGCAATGCTCAGAGCG 710
DB 1728 GTGATGTCTCTAGTGTGCGGCC 1752
```

RESULT 15

AA012552
ID AA012552 standard; DNA; 2756 BP.

XX AA012552;

XX 23-SEP-1991 (first entry)

XX Encodes plasminogen mutcin T13 with thrombin cleavage site.

XX protease; fibrinolysis; blood clotting; ss.

XX Key Location/Qualifiers

FT mutation 1796..1804

FT /tag= a

FT /note= "replaces CCTGGA (i.e. two codons)"

```
PN W09109118-A.
XX 27-JUN-1991.
PD 07-DEC-1990; 90WO-G001912.
XX 07-DEC-1989; 89GB-0027722.
PR 07-DEC-1990; 90WO-GB01911.
XX (BRB-) BRIT BIO-TECHN LTD.
XX Dawson KM, Edwards RM, Forman JM;
XX WPI, 1991-208145/28.
XX P-PDB; AARI2943.
PS Activatable fibrinolytic and antithrombic proteins - activated by
XX e.g. factor Xa, thrombin or activated protein C
XX Claim 16; Fig 2 and Fig 5; 73pp; English.
XX This sequence encodes a plasminogen mutant cleavable by thrombin.
XX Activation is localised to the thrombus because cleavage to plasmin
XX is by an enzyme of the blood clotting pathway. Compositions
XX comprising the mutant plasminogen encoded by this sequence are used
XX for treatment or prevention of thrombosis, etc.
XX See AA012542-Q12558.
XX Sequence 2756 BP; 765 A; 671 C; 695 G; 625 T; 0 other;
OY 92 CAGCAGCAGCTGACGAAAGCAGTGGTCCAGATGCTACCAATGGTGTGACAGAGTT 151
DB 1138 CAGCAGCAGCTGACGAAAGCAGTGGTCCAGATGCTACCAATGGTGTGACAGAGTT 1217
OY 152 ATCGAGCAGCTGCTCCAGCAGCTGTCCAGAGAGAGCTCCCAAGCTGTGATCATGA 211
DB 1218 ACCGAGGACATCTCCACACCCACACAGGAAAGAGTGTCACTGTGATCATGA 1277
OY 212 CACCATCAACATATATATAGGACACAGAAATACCAATGGTGTGATCATGA 271
DB 1278 CACCATCAACATATATATAGGACACAGAAATACCAATGGTGTGATCATGA 1337
OY 272 ACTGAGGAAATCAGATGCTGTGAGCAGCTCTTATTTATACGAGGAGTCCGGGTGCA 331
DB 1338 ACTGAGGAAATCAGATGCTGTGAGCAGCTCTTATTTATACGAGGAGTCCGGGTGCA 1397
OY 332 GGTGGAGTACTGCAACCTGACGCAATGCTCAGACAGCAGAGGAGTCCGCGGCTC 391
DB 1398 GGTGGAGTACTGCAACCTGACGCAATGCTCAGACAGCAGAGGAGTCCGCGGCTC 1457
OY 392 CGACGTGATACCCGCTTCCAGAGCTGCTCCGAAACAGCAGCAGTGAAGCA 451
DB 1458 CGCCTGTGCTCTGCTCCAGATGTAGAGTCTCTCCGAAAGCA 1501
OY 452 GGCCTGGGTGACGAGTGTCTACCATGTATATGACAGAGTATTCAGGCAATCTCA 511
DB 1502 -----GACTGTATGTTTGGAGATGGGAAAGATACCGAGCAAGAGGCGCA 1547
OY 512 CCAGTGTACAGGAAAGCCTGCCAAGCTTGATCATATGACACAGCAGTGCATAG-- 569
DB 1548 CCAGTGTACAGGAAAGCCTGCCAAGCTTGATCATATGACACAGCAGTGCATAG-- 1607
OY 570 -TCGAGCCCAAGATATCTACCAATGCTGTGATCATGATGATCTGAGAGATCCAG 628
DB 1608 TTTTACCTCCAGAGCAATTCAGGAGGCTGTGAAAAAAATTTACTGCGTAACTCTG 1667
OY 629 A---TGCTGTGAGCTCCTTATTTATATGAGAGGATCCCGGTGTCAGTGGAGTACT 685
DB 1668 ATGCTGTGAGCTCCTTATTTATATGAGAGGATCCCGGTGTCAGTGGAGTACT 1727
```

Oy 686 GCACCTGACGCAATGCTCAGACGC 710
| | | | | | | | | |
Db 1728 GTGATGTCCCTCAGTGTGCGGCCCC 1752

Search completed: March 6, 2003, 05:12:19
Job time : 932 secs

. . .

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 06:19:18 ; Search time 148 Seconds
(without alignments)
10360.701 Million cell updates/sec

Title: US-09-923-515-3_COPY_1_5000

Sequence score: 5000
1 ctggagattggagacactt.....caagcttgatcatatgac 5000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	236.6	4.7	2497	1	US-08-643-219-12
2	236.6	4.7	2497	2	US-09-131-995-12
3	236.6	4.7	2497	2	US-08-832-087B-12
4	236.6	4.7	2497	3	US-08-851-350-12
5	236.6	4.7	2497	4	US-09-132-154-12
6	236.6	4.7	2679	6	5200340-7
7	236.6	4.7	2753	1	US-07-854-603-1
8	235	4.7	2296	3	US-07-750-080A-18
9	235	4.7	2296	3	US-08-651-472-18
10	235	4.7	2296	3	US-08-358-928-18
11	216.2	4.3	1134	4	US-09-206-059-29
12	92.4	1.8	2542	1	US-08-441-370-1
13	67.6	1.4	645	3	US-08-985-526-22
14	67.6	1.4	1284	3	US-08-985-526-24
15	62.8	1.3	2172	4	US-08-030-410-2
16	62.8	1.3	2184	4	US-07-815-333A-1
17	62.8	1.3	2288	1	US-08-290-937B-4
18	62.8	1.3	2289	1	US-07-838-410-2
19	59.2	1.2	2188	1	US-07-882-925A-4
20	59.2	1.2	2188	1	US-08-184-012C-4
21	55.2	1.1	315	1	US-08-243-035A-11
22	54.8	1.1	2021	1	US-07-882-925A-3
23	54.8	1.1	2021	1	US-08-184-012C-3
24	54.8	1.1	2216	2	US-08-666-082B-2
25	54.8	1.1	2219	1	US-07-883-925A-1
26	54.8	1.1	2219	1	US-07-883-925A-2
27	54.8	1.1	2219	1	US-08-184-012C-1

28	54.8	1.1	2219	1	US-08-184-012C-2	Sequence 2, Appl1
29	54.8	1.1	2232	1	US-08-334-177-1	Sequence 1, Appl1
30	54.8	1.1	2232	5	PCT-US95-13830-1	Sequence 1, Appl1
31	54.8	1.1	2262	1	US-07-882-925A-7	Sequence 7, Appl1
32	54.8	1.1	2262	1	US-08-184-012C-7	Sequence 7, Appl1
33	51.4	1.0	3398	5	PCT-US95-08493-12	Sequence 12, Appl1
34	35.6	0.7	6100	1	US-07-882-925A-6	Sequence 6, Appl1
35	35.6	0.7	6100	1	US-08-184-012C-6	Sequence 6, Appl1
36	33.4	0.7	6751	1	US-07-882-925A-5	Sequence 5, Appl1
37	33.4	0.7	6751	1	US-08-184-012C-5	Sequence 5, Appl1
38	32.6	0.7	4092	2	US-08-469-537A-106	Sequence 106, Appl
39	32.2	0.6	705	4	US-08-998-416-1136	Sequence 1136, Ap
40	32	0.6	32	2	US-07-832-905B-19	Sequence 19, Appl
41	32	0.6	32	2	US-08-700-737-19	Sequence 19, Appl
42	31.8	0.6	4626	1	US-08-306-691B-22	Sequence 22, Appl
43	31.8	0.6	4626	5	PCT-US93-06251-27	Sequence 27, Appl
44	31.6	0.6	1869	3	US-08-952-967-7	Sequence 7, Appl1
45	30.8	0.6	2058	4	US-09-560-639-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-08-643-219-12
; Sequence 12, Application US/08643219
; Patent No. 5801146
; GENERAL INFORMATION:
; APPLICANT: Davidson, Donald J.
; TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES
; TITLE OF INVENTION: AND METHODS FOR INHIBITING ANGIOGENESIS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/643, 219
; FILING DATE: 06-MAY-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Casuto, Dianne
; REGISTRATION NUMBER: 40,943
; REFERENCE/DOCKET NUMBER: 5940.US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-938-3137
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2497 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-643-219-12

Query Match 4.7%; Score 236.6; DB 1; Length 2497;
Best local Similarity 58.6%; Pred.No. 1.3e-70;
Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;
QY 809 AGTGTACATGTATGACAGAGTATCGAGGCACATCTCCACCACTGTCTACAGAA 866

```

Db      870 AGTGTCTGAGGGAACAGGTGAAACATATCGCGGGAATGmGGCTTAACTGCTCGGGGC 929
      869 GAACCTGCCAAGCTTGGTCACTATGACACCACTGGCTAGTGGAGCCCAAGTAAT 928
      930 ACACCTGTAGAGTGGAGTGCACAGACCCCTGCACACATTAACAGAACACCGAAACT 989
      929 ACCCAATGCTGGCTTGCATGTAACACTGACAGAAATCCAGATGCTGGACGCTCTT 988
      990 TCCCTGGAAAAATTTGGATGAAAACTACGCCCAATCCTGACGGAAAAAGGCCCAT 1049
      989 ATTGTATATACAGAGGATCCCGGTGCAGGTGGAGTACTCAACCTGACGCAATGCTCAG 1048
      1050 GGTGCAATACCAACCAACGCCCAATGGGTGGAGTACTTAAGTAACGCT----- 1100
      1049 ACGCAGAAAGGACTGCGGCTCGGCTGACCTGTTACCCGGTTCAGAGCTAGAGGCTC 1108
      1101 -----CCTGTGACTCTCTCCCAAGTATCCACG-----GAACAT 1133
      1109 CTTCGCAACAAAGCAGCAGTACGCAAAAGCCCTGGGGTGCAGAGTGTCACTATGTAATG 1168
      1134 TGGCTCCACAGACACACACTGTAGGTAAACCTGTGGTCCAGAGCTCACTACATGTGATG 1193
      1169 GACAGATTATGAGGACATATCTCCACACTGTCAAGAGAAAGACCTGCAAGCTTGGT 1228
      1194 GACAGAGCTACCGAGGACATCTCTCAACACACAGGAAAGAAAGTGTGCTTGGT 1253
      1229 CATCTATACACACACATCGCATGTGGACCCCAAGTAATCTAACCAATGCTGCTTGA 1288
      1234 CATCTATACACACACACCGGACCCAGAAAGCCCAAGTAATCTAACCAATGCTGCTTGA 1313
      1289 TCATGAACTACTGAGAAATCAGATCTGTGACGCTCTTATTTATGAGAGGATC 1348
      1314 CAATGAACTACTGAGAGGATCCAGATGCGATTAAGGCCCTGTGTTTACACAGATC 1373
      1349 CCGGTGTAGTGGAGTACTGCAACTGAGCAATGCTCAGACGAGAAAGGAGTCCG 1408
      1374 CAGGCTGAGGTGAGATGCACTGAACTGAAATGCTCAGAGAAAGAGAGTGTG 1433
      1409 TCGGCTCTCAGACTGTACCCGGGTTCAGAGCTAGAGGCTCTTCCGAAACAGACGA 1468
      1434 TAGCACTCTCGGCTGTGTCTCTGCTTCAGATGTAAGAGACTCTTCCGAAAG----- 1486
      1469 CTGACAAAGGCTGGGGTGCAGAGTCAATGTAATGAGACAGATTAATGAGGCA 1528
      1487 -----GACTGATGTTTGGGAAATGGGAAAGATACCGAGCA 1523
      1529 CATCTCCACACTGTACAGAGAAACCTGCCAAGCTGGTCACTATGACACCAACT 1588
      1524 AGAGGGCAGACACTGTATGAGAGCCATGCCAGACTGGCTGCGCCAGAGAGCCCAT 1583
      1589 GCGATAG--TGGAGCCCAAGATATCAACCAATGCTGGCTTGAATGATGACTGTGCA 1645
      1584 GACACAGCATTTTCACTCCAGAGACAAATCCACGGGGGCTGGAAAAAATTTACTGCC 1643
      1646 GGAATCCAGA---TGCGTGGCAGCTCTTATGTAATAGAGAGATCCCGGTGTCAGGT 1702
      1644 GTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1703
      1703 GGGAGTACTGCAACCTGACGCAATGCTCAGAGCG 1736
      1704 ACGACTACTGTGATGTCCCTCAGTGTGCGGCC 1737

```

RESULT 2
US-09-131-995-12
Sequence 12, Application US/09131995

Patent No. 5972896
GENERAL INFORMATION:

APPLICANT: Davidson, Donald J.
TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES

AND METHODS FOR INHIBITING ANGIOGENESIS
NUMBER OF SEQUENCES: 14

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/131,995
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/832,087
FILING DATE: 03-APR-1997
APPLICATION NUMBER: 08/643,219
FILING DATE: 06-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Casuto, Dianne
REGISTRATION NUMBER: 40,943
REFERENCE/DOCKET NUMBER: 5940.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-938-3137
TELEFAX: 847-938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2497 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-131-995-12

```

Query Match 4.7%; Score 236.6; DB 2; Length 2497;

Best Local Similarity 58.6%; Pred. No. 1.3e-70; Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;

```

      809 AGTGTCAATGTAATGAGACAGATTAATGAGCACTACTCCACCACTGTGACAGAA 868
      870 AGTGTCTGAGGGAACAGGTGAAACATATCGCGGGAATGCTGTTACCGTCCGGGC 929
      869 GAACCTGCCAAGCTTGGTCACTATGACACCACTGCAAGTGTGAGACCCCAAGTAAT 928
      930 ACACCTGTAGAGTGGAGTGCACAGACCCCTGCACACATTAACAGAGAACGAAAACT 989
      929 ACCCAATGCTGGCTTGAATGATGAACTACTGCAAGAAATCCAGATGCTGTGGACGCTCTT 988
      990 TCCCTGGAAAAATTTGGATGAAAACTACTGCCCAATCTGAGAGGAAAAAGGCCCAT 1049
      989 ATTGTATATACAGAGGATCCCGGTGCAGTGGAGTACTGCAACCTGACCAATGCTCAG 1048
      1050 GGTGCAATACCAACCAACAGCAAGTGGGTGGAGTACTGTAAGTACCT----- 1100
      1049 ACGCAGAAAGGACTGCGGCTCGGCTCCGACTGTTACCCGGTTCAGAGCTAGAGGCTC 1108
      1101 -----CCTGTGACTCTCTCCCAAGTATCCACG-----GAACAT 1133
      1109 CTTCGCAACAAAGCAGCAGTACGCAAAAGCCCTGGGGTGCAGAGTGTCACTATGTAATG 1168
      1134 TGGCTCCACAGACACACACTGTAGGTAAACCTGTGGTCCAGAGCTCACTACATGTGATG 1193
      1169 GACAGATTATGAGGACATATCTCCACACTGTCAAGAGAAAGACCTGCAAGCTTGGT 1228
      1194 GACAGAGCTACCGAGGACATCTCTCAACACACAGGAAAGAAAGTGTGCTTGGT 1253
      1229 CATCTATGACACACACTGCGATGATGTCGAGCCCAAGTAATCTAACCAATGCTGCTTGA 1288
      1234 CATCTATGACACACACCGGACCCAGAGAGCCCAAGAAATCTAACCAATGCTGCTTGA 1313

```


ATTORNEY/AGENT INFORMATION:

NAME: Casuto, Dianne
REGISTRATION NUMBER:

REGISTRATION NUMBER: 40,943
REFERENCE/DOCKET NUMBER: 59

TELECOMMUNICATION INFORMATION:

TELEPHONE: 847-938-3137

TELEFAX: 84 / -938-2623
TELEX:

INFORMATION FOR SEQ ID NO:

SEQUENCE CHARACTERISTICS:

LENGTH: 249 / base pairs
TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA
9-133-15A-12

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

ery Match

Local Similarity

04/1/

809 AGT

Lane 1
Lane 2
Lane 3
Lane 4
Lane 5
Lane 6
Lane 7
Lane 8
Lane 9
Lane 10
Lane 11
Lane 12

8/0 AGTGTCTGAAGGGAACACAGGTGAATACTATCCGGGAATGTGGCTGTACCGTGTCCGGCC 929

QY	865	GAACCTGCCAAGCTTGGTCAATATGACACCACTGCGATAGTCGAGCCCGAGAACTACT	928
Db	930	ACACCTGTGCACACTGGAAGTGCACACACCCCTTCACACACATTAACAGGACACCGAAAACT	989
QY	929	ACCCAAAATGCTGGCTTATATATGAACTATGCGAGAAATCCAGATGCTGTGTGGACGCTCTT	988
Db	990	TCCCTCGCAAAATATTTGGATGAAATACTATGCGGCAATCTGAGGAGAAAAAGGCCCAT	1049
QY	989	ATTGTTATACGAGGAAATCCCGGTGTGATGAGTGGAGATATCTGCAACTGACCAATGCTGC	1048
Db	1050	GGTGGCATACACCAACAAGCCAAAGTGGGTGGGATACGTGAATATACCTG-----	11000
QY	1049	ACCGAAGAGGAACTGCCGTGCGCCTTCGACACTGTTACCCGGTTCCAAAGCTTAGAGGCTC	1108
Db	1101	-----CGTGTATCTCCCTCCCAATATCCAGC---GAACAAT	1133
QY	1109	CTTCCGAACAAGACCGCACTGAGCAAAAGGCGTGGGGTGCAGAGTGTCTACATGCTAATG	1168
Db	1134	TGGCTCCCAAGACACCACTGAGACTAAACCCGTGTGTCCAGAGATGCTACCAATGGTATG	1193
QY	1169	GACAGAGTTATCGAGGAGACATCTCCACACAGTGTGCACAGGAAGAACTGTCCAAAGCTTGT	1228
Db	1194	GACACAGCTACCGAGGCACATCTCCACACACACACACAGGAAGAAATGTCACTTGTGT	1255
QY	1229	CATCTATGACACCACTGCGATATGTGGAGCCCGAATATCAACCAATGCTGGCTTGA	1288
Db	1254	CATCTATGACACCACTGCGGACCAAGAACCCCGAATAATCACTCCAAATGCTGGCTTGA	1313
QY	1289	TCATGAACTACAGAGAAATCCAGATGCTGTGGAGCTCTTATTTGTTATACGAGGAATC	1348
Db	1314	CATGAACTACTGAGAGAAATCCAGATGCCGATAAAGGCCCTGTGTTTATCCACAGAAC	1373
QY	1349	CCGCTGTCAAGTGGGAGTACTGCCAACCTTAGCCAAATGCTCAGACCGAGGAAGGACTGCCG	1408
Db	1374	CCAGGCTGAGGTGGGAATACTGCAACCTTAAAAAATGCTCAGGAACGAAAGCAGAGTGGT	1433
QY	1409	TCGCGCTCCAGACTGTTACCCCGTGTCCAAAGCTTAGAGGCTCTCTCCGAACAAGCACGA	1468
Db	1434	TAGCACTCCCGCTGTTGTCTCTTCCTCCAGATGTAGAGACTCTCTCCGAAGAA-----	1488
QY	1469	CTGAGCAAGAGGCTGGGGGTGCAGAGATGCTACCAATGTAATAGACAAAGTATTACGAGCA	1528
Db	1487	-----GACTGTATGTTTGGGAATGGGAAGGATACCGAGGGA	1522
QY	1529	CATATCCACACACTGTGCACAGGAAGCAACTGCCAAAGCTTGGTCACTTATATACACCACT	1588
Db	1524	AGAGGGCCACCACTGTACTGTGGAGCCCACTGCACAGATGGGGCTGCCAAGAGCCCAATA	158

[illegible]

RESULT 6

5200340-7
Patent No. 5200340

APPLICANT: FOS

;PATRICK J.;PINGEL, KURT;YOSHITAKE, SHINJI

TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
ACTIVATORS

```

;CLLIVATIONS
: NUMBER OF SEQUENCES: 34

```

CURRENT APPLICATION DATA

APPLICATION NUMBER:

FILING DATE: 22-MAY-1987
; ;
; SEO ID NO.: 7;

```

;SEQ ID NO: 1
;
LENGTH: 2679

```

5200340-7

Quarter Match

Best Local Similarity

Matche

900 ACFT

009 AGIGTACCAIGGTAATGGACAGATTATCGAGGCACATATCTCCACCACTGTCAC

Db 842 AGTGTCTGAAGGACAGGTGAAACTATCGCGGGAATGTGGCTGTTACCGTGTCCGGGC 901

QY	869	GAACTGCGAAGCTTGGTCATCTATGACACACACACTGCGATAGTGGAGCCGACATACT	928
Db	902	ACACTGTACACACTGGAGGTGCACAGACCCTTCACACATTAACAGACACGAAACT	961
QY	929	ACCCAAATGCTGGGTGATCATGAACTACTCGCGAAGAACCCAGTGTGTGGAGAGTCCCT	988
Db	962	TCCCTTCGCAAAATTTGGATGAAATCTACGCTCCGCAATCTGTACGGAAGAAAGGCCCAT	1021
QY	989	ATTGTTATACGAGGAGATCCCGGTGTCAGTGGGAGTACTGCAACTGACGCATCTCAG	1048
Db	1022	GGTGCCATACCAACCAACAGCCAAATGGGTGGGAGTACTGTAAATACCGT-----	1072
QY	1049	ACGCGAAGGAGACTGCGCTGCGCGCTCGAGTCTGTACCCCGGTTCACACCTAAGAGGCTC	1108
Db	1073	-----CCTGTACTCTCTCCCACTATTCACG-----GAACAAT	1105
QY	1109	CTTCCGAACACACCCGACTGAGCAAAAGCCTGGGGTCAGAGTGTCTACATGCTAATG	1168
Db	1106	TGGCTCCCAACAGCACACCTGAGTAACTCCCTGTGGTCCAGGACTCTCATGTGTTAGT	1167
QY	1169	GACACAGTTATCGAGGCGACATATCTCCACACTGTGCACAGGAACAACTCCAGACTGT	1228
Db	1166	GACACAGCTTACCGAGGACATCTCTCCACACACACACAGGAAGAAAGTTCAGTCTTGGT	1225
QY	1229	CATCTATACACCACTGCGATATGTGGACCCAGATATCTACCAATGCTGGCTTGA	1288
Db	1226	CATCTATGACACACACCGGCGACCAAGAAACCCAGAAATCTACCAATGTGTGCTCGA	1285
QY	1289	TCAATCACTACTGAGAGATCCAGATCTCTGGCAGCTCTCTATTGTTTATACGAGGATC	1348
Db	1286	CAATCACTACTGCGAGGATCCAAATCCGATTAAGAGCCCTCGTGTGTTTATACACAGAC	1345
QY	1349	CCGCTGTAGAGTGGGAGAGTACGCAACTGACAGCAATGCTTAGACACGGAAGGACGTGCG	1408
Db	1346	CCAGGCTTAGAGTGGGAGTACTGCAACTCGAAAAAATGTCTGAGACACGAAAGCAAGTGTGG	1405
QY	1409	TGCGCCTCCGACTGTATCCCGGTTCCAAAGCTAGAGGCTCTCTTCGAAACAGCACCA	1468

```
Db 1406 TAGCACCCTCGCCTGTTCCTGCTGCAGATGTAGAGACCTCTCCGAGAA----- 1458
QY 1469 CTGAGCAAGGCGCTGGGGTGCAGAGTCTACCATGTGATGACAGAGTTATCGAGCA 1528
Db 1459 -----GACTGTATGTTTGGGAATGGAAAGGATACCGAGCA 1495
QY 1529 CATACTCCACCATCTGCACAGGAAGACCTGCAAGCTTGTCATCTATGACACACACT 1588
Db 1496 AGAGGCGGACACCTGTTACTGAGGAGCGCATGCCAGACTGGGGTCCGACAGAGCCCCATA 1555
QY 1589 CCGATAG---TCGGACCCCAATATCTACCAATGCTGGCTGATCATCACTACTGCA 1645
Db 1556 GACACAGCATTTTCACTCCAGAGCAAAATCCAGCGGGCTGCGAAAAAAATTTACTGCG 1615
QY 1646 GGAATCCAGA---TGCTGTGCGACCTCTTATGTATAGAGAGGATCCCGGTGTCAGGT 1702
Db 1616 GTACCCGTGATGTGATGTAGTGTGCTGCTGCTACAGACAAATCCAGAAACTTT 1675
QY 1703 GGGAGTACTGCACACTGACCGCAATGCTCAGACGC 1736
Db 1676 ACGACTACTGTGATGTCCCTCAGTGTGGGCC 1709
```

RESULT 7

```
US-07-854-603-1
; Sequence 1, Application US/07854603
; Patent No. 5637492
```

GENERAL INFORMATION:

```
APPLICANT: Dawson, Keith M
APPLICANT: Edwards, Richard M
APPLICANT: Forman, Joan M
TITLE OF INVENTION: Activatable fibrinolytic and
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESSES:
ADDRESSER: Dr. John J. McDonnell
STREET: Ten South Wacker Drive, Suite 3000
CITY: Chicago
STATE: IL
COUNTRY: USA
```

COMPUTER READABLE FORM:

```
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/854, 603
FILING DATE: 19901207
```

CLASSIFICATION:

```
ATTORNEY/AGENT INFORMATION:
NAME: McDonnell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,338
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
```

INFORMATION FOR SEQ ID NO: 1:

```
SEQUENCE CHARACTERISTICS:
LENGTH: 2753 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
```

FEATURE:

```
NAME/KEY: misc_feature
LOCATION: 1..2753
OTHER INFORMATION: /note="Fig. 2 Plasminogen cDNA
FEATURE:
NAME/KEY: sig_peptide
```

```
LOCATION: 65..121
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 122..2494
FEATURE:
NAME/KEY: CDS
LOCATION: 65..2494
FEATURE:
NAME/KEY: misc_feature
LOCATION: 54..55
OTHER INFORMATION: /note="Ball site"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2564..2565
OTHER INFORMATION: /note="Spl site"
US-07-854-603-1
```

Query Match

Best Local Similarity 4.7%; Score 236.6; DB 1; Length 2753;

Matches 410; Conservative 0; Mismatches 179; Indels 36; Gaps 3;

```
QY 92 CACGACGACCTGAGCAAGCATGTGTGTCAGAGATTGCTACCATGTGATGACAGAGTT 151
Db 1158 CAGACACACCTGAGTAACTCCCTGTGTGTCAGAGACTGCTACCATGTGATGACAGAGTT 1217
QY 152 ATGAGGACGCTACTGCTACCATGTCACAGAGAGAGACCTGCCAAGCTTGGTCATATGA 211
Db 1218 ACCGAGGACATCTCTCCACACACACAGAGAAAGTGTCACTTGTGTCATATGA 1277
QY 212 CACCATCAACATANTAGAGACACAGAAACTACCCAAATGCTGGCTTGATGATGAAT 271
Db 1278 CACCATCAACATANTAGAGACACAGAAACTACCCAAATGCTGGCTTGATGATGAAT 1337
QY 272 ACTGACGAGATCCAGATGCTGTGAGAGCTCTTATGTATAGAGAGATCCGGGTGTA 331
Db 1338 ACTGACGAGATCCAGATGCTGTGAGAGCTCTTATGTATAGAGAGATCCGGGTGTA 1397
QY 332 GGTGGAGTACTGCAACCTGACGCAATGCTCAGACGCGAAGGAGCTGCGCGCTC 391
Db 1398 GGTGGAGTACTGCAACCTGACGCAATGCTCAGACGCGAAGGAGCTGCGCGCTC 1457
QY 392 CGACTGTACCCCGGTCCAAAGCTAGAGCTCCTCCGAAACAGACCCAGTACGAGCA 451
Db 1458 CGCTGTGTCTGCTCTCCAGATGTAGAGACTCTCCGAGAA----- 1501
QY 452 GGCCTGGGGTGCAGAGTGTCTACATGTTAATGACAGAGTTATGAGGACATACTCA 511
Db 1502 -----GACTGTATGTTTGGGAATGGAAAGGATACCGAGGACGAGGCGCA 1547
QY 512 CCATGTCACAGAGAAAGCTGCAAGCTGTGTCATCTATGACACCATCTCCATAG-- 569
Db 1548 CCATGTCACAGAGAAAGCTGCAAGCTGTGTCATCTATGACACCATCTCCATAG-- 1607
QY 570 -TCGACCCCAAGATACTACCCAAATGCTGCTTGTATCATGAACTACTGAGGAATCCAG 628
Db 1608 TTTTCATCCCAAGATACTACCCAAATGCTGCGGGGTGTGAAAAAATTAAGTCCATACCTG 1667
QY 629 A---TCGCTGCGACCTCTTATGTTATAGAGAGATCCCGGTGTAGTGGAGTACT 685
Db 1668 ATGTATGTATGATGTGCTGCTGTCTACAGCAAAATCCAGAAAACTTTAGACTACT 1727
QY 686 GCAACCTGACGCAATGCTCAGACGC 710
Db 1728 GTGATGTCCCTCAGTGTGGGCC 1752
```

RESULT 8

```
US-07-750-080A-18
; Sequence 18, Application US/07750080A
; Patent No. 5445953
```

GENERAL INFORMATION:

```
APPLICANT: DORNER, F.
APPLICANT: SCHERFLINGER, F.
```


REFERENCE/DOCKET NUMBER: 30472/166/1MMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2296 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid;
DESCRIPTION: Synthetic DNA oligonucleotide
IMMEDIATE SOURCE:
CLONE: pN29pt-LP9
US-08-651-472-18

Query Match 4.7%: Score 235; DB 3; Length 2296;
Best Local Similarity 58.5%; Pred. No. 4.2e-70;
Matches 546; Conservative 0; Mismatches 316; Indels 72; Gaps 5;

QY 467 AGTGTACCATGTAATGAGAGATTATGAGGCAATATCTCACAGGAA 526
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 610 AGTGTCTGAGGAGACAGTGAATACTACGCGGAAATGCGCTTACGTTCCGGGC 669
QY 527 GAACGCGCAAGCTGTGTCATCTATGACACCACTGCGATAGTGGACCCGCAATACT 586
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 670 AACCTGTACAGCTGAGAGTGACAGACCCCTCACACATTAACAGACACCAAAACT 729
QY 587 ACCCAATGCTGCTGATCATGAACTACTGACAGAAATCCAGATGCTGGCAGCTCTT 646
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 730 TCCCTGCAAAAATTTGATGATAAACTACGCGCAATCCTGACGAAAGGCGCCCAT 789
QY 647 ATTGTATATGAGAGATCCGCTGTCAGTGAGTGAGTCTGCACTGACATGCTCATG 706
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 790 GGTGCAATACCAACCAAGCAAGTGGGTGAGTCTGTAAGATACCGT----- 840
QY 707 AGCGAAGAGGACTCGCTCGGCTCGAGCTGTACCCCGTTCCAGGCTTAGAGGCTC 766
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 841 -----CCTGTGACTCTCTCCCGATGTCACG---GAACAAAT 873
QY 767 CTTCGAAACAGACACGATGAGCAAAAGGCTGGGGTGTGAGAGTCTACATGTAATG 826
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 874 TGGCTCCACAGCACACCTGAGCTAACCCCTGTGTCAGAGCTCTACACAGGTGAG 933
QY 827 GACAGAGTATGAGGAGCAGATACACACACTGTCAACAGAAACCTGCCAAGCTGGT 886
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 934 GACAGAGCTACCGAGGACATCTCTCACACACACAGAAAGAGTGTCACTTGGT 993
QY 887 CATCTATGACACACACTCGATATGTGAGACCCCAATATCTACCAATGCTGGCTTGA 946
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 994 CATCTATGACACACACCGGACCAAGAAAGCCCAAAATCTACCAATGCTGGCTTGA 1053
QY 947 TCATGAACTACGAGGAATCCAGATGCTGTGAGAGTCTTATTTATGAGAGGATG 1006
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 1054 CATGAACTACGAGGAATCCAGATGCTGTGAGAGTCTTATTTATGAGAGGATG 1113
QY 1007 CCGGTGTAGGTGAGGAGTACTGCAACCTGACAGCAATCTCAGACGAGAGGAGTCCG 1066
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 1114 CAGGCTCAGGTGAGGAGTACTGCAACCTGAAATAATGCTCAGAGACAGAGGTGG 1173
QY 1067 TGGCGCTCCGAGCTTTCACCGGCTGCAAGGCTAGAGGCTCTCTCCGAAACGACCA 1126
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 1174 TACGACTCTCCGCTGTCTCTGCTCCAGATGTAGAGACTCTCTCCGAAAGAA----- 1226
QY 1127 CTGAGCAAGAGCTGGGTGAGAGTGTACATGTGTAATGACAGAGTATTCAGAGCA 1186
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 1227 -----GACTGTATGTTTGGGAATGGAAAGATACCGAGCA 1263
QY 1187 CATATCCACACATGTCACAGAGAAACCTGCAAGCTTGGTATCTATGACACCACT 1246
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 1264 AGAGGCGACACACTGTCTACTGTGAGAGCCATGCCAGAGTGTGCCACAGAGCCCATTA 1323

QY 1247 CGCATAG---TCGACCCCAAGAACTACTACCAATGCTGCTGTGATCATGAATCTGCA 1303
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 1324 GACACGCAATTTTACATCCAGACAAATCCACGGGGGTCTGAAAAAATTTACTGCC 1383
QY 1304 GGAATCCAGA--TCTGTGACAGCTCTCTTATTTATACAGAGATCCCGGTGTCAGGT 1360
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 1384 GTAACCTGATGATGATAGTGTGCTCCGTGTGCTACACGACAAATCCAGAAACTTT 1443
QY 1361 GGGAGTACGACACTGACGCAATGCTCAGACCG 1394
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 1444 ACAGTACTGTGATGTCTCCCTCAGTGTGCGGCC 1477

RESULT 10
US-08-358-928-18
Sequence 18, Application US/08358928
Patent No. 6265183

GENERAL INFORMATION:

APPLICANT: DORNER, Friedrich
APPLICANT: SCHEFFLINGER, Friedrich
APPLICANT: FALKNER, Falko Gunter
APPLICANT: PLEIDERER, Michael
TITLE OF INVENTION: DIRECT MOLECULAR CLONING OF CHIMERIC
TITLE OF INVENTION: VIRUSES CONTAINING HUMAN IMMUNODEFICIENCY VIRUS TYPE 1
TITLE OF INVENTION: (HIV-1) ANTIGENS
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Hardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/358,928
FILING DATE:

FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/914,738
FILING DATE: 20-JUL-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/750,080
FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/166/1MMU

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2296 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid;
DESCRIPTION: Synthetic DNA oligonucleotide
IMMEDIATE SOURCE:
CLONE: pN29pt-LP9
US-08-358-928-18

Query Match 4.7%: Score 235; DB 4; Length 2296;
Best Local Similarity 58.5%; Pred. No. 4.2e-70;
Matches 546; Conservative 0; Mismatches 316; Indels 72; Gaps 5;

QY 467 AGTGTACCATGTAATGAGAGATTATGAGGCAATATCTCACAGGAA 526

```
Db 610 AGTGTCTGAAAGGAGACAGAGTGAAGAACTATCGGGGAAATGTGGCTGTACCTGTTCCGGGGC 669
      |||||
Qy 527 GAACCTGCAAGCTTGGTCATCTANGACACACACTCCGATAGTCSGACCCCAAGAACT 586
      |||||
Db 670 ACACCTGACACACTGGAGTGCACAGACCCCTCACACACATACAGAGACACCAAGAAACT 729
      |||||
Qy 587 ACCCAATGCTGGCTTGATCATGACTACTGAGAAATCCAGATGCTGGAGCTCTT 646
      |||||
Db 730 TCCCTGCAAAATTTGATGAAATACTACTGCGCAATCTGAGAGGAAAAAGGCCCAT 789
      |||||
Qy 647 ATTGTATACAGAGGATCCGGTGTGAGTGGAGTACTGCAACCTGACCCATGCTCAG 706
      |||||
Db 790 GGTGTCATACAAACCAACAGCAAGTGGTGGATGATGTAATACCTG----- 840
      |||||
Qy 707 ACGAGAAAGGACTGCGTGCCTCCGACTGTTACCCCGTTCGAAGCTTACAGGCTC 766
      |||||
Db 841 -----CTGTGACTCTCCCAATATCCAGC-----GAACAAAT 873
      |||||
Qy 767 CTTCGAAACAGACACGACTGAGCAAGGCTGGGGTGGAGAGTGTACCATGTGTAATG 826
      |||||
Db 874 TGGCTCCACAGACACACCTGACTTAACCTGTGTGCTCAGAGACTGCTACGAGGTGATG 933
      |||||
Qy 827 GACAGAGTTATCGAGGACATCTCCACACTGTGCACAGAGAAACCTGCCAAGCTTGT 886
      |||||
Db 934 GACAGAGCTACCGAGGACATCTCCACACACACAGAGAAAGTGTAGTCTTGTGT 993
      |||||
Qy 887 CATCATGACACACACTGCTAGTGCAGACCCCAAGAAATGCTACCCAAATGCTGAGCTGA 946
      |||||
Db 994 CATCATGACACACACCGGACACAGAAAGACCCCAAGAAATGCTGAGCTGA 1053
      |||||
Qy 947 TCATCAATGACTGAGAGAAATCCAGATGCTGAGAGCTCTTATGTTATACAGAGGATC 1006
      |||||
Db 1054 CATCAATGACTGAGAGAAATCCAGATGCTGAGAGCTCTTATGTTATACAGAGATC 1113
      |||||
Qy 1007 CCGGTGTGAGTGGAGTACTGCAACCTGACAGCAATGCTCAGAGAGGAGGACTGCCG 1066
      |||||
Db 1114 CCAGGCTGAGTGGAGTACTGCAACCTGAAAAAATGCTCAGAGAAACGAGTGTG 1173
      |||||
Qy 1067 TCGCGCTTCGACTGTACCGCGTTCCAAAGGCTTGAAGGCTTCCGAACCAACACCGA 1126
      |||||
Db 1174 TTAGCACTCCGCGCTGTGTCTCTTCAGATGTAGAGATCTCTCCGAAGAA----- 1226
      |||||
Qy 1127 CTGAGCAAAAGGCTGGGGTGCAGAGATGCTACCATGTGTAAAGAGAGATTTAGAGGCA 1186
      |||||
Db 1227 -----GACTGTATGTTTGGAAATGGGAAAGATACCGAGGCA 1263
      |||||
Qy 1187 CATACTCCACACTGTGCACAGAAAGCTGCCAAGCTTGTGTCATGTATGACACACACT 1246
      |||||
Db 1264 AGAGGGCGACACTGTACTGGAGCGCAGCAGACTGGGCTGCCAGAGGCCCATTA 1323
      |||||
Qy 1247 GCGATAG---TCGAGCCCGAATGACTACCAAAATGCTGGCTGTATCTGAACTAGTGA 1303
      |||||
Db 1324 GACACAGCAATTTTCACTTCACAGACAAATTCACAGGCGGGTCTGAAAAAATTTACTGCC 1383
      |||||
Qy 1304 GGAATCCAGA---TGCTGTGCAAGCTCTTATTGTTATACGAGGATCCCGTCTCAGGT 1360
      |||||
Db 1384 GTAAACCGATGTGATGATGAGTGGTCCCTGGTGTCTACACGCAAAATCCAGAAACTTT 1443
      |||||
Qy 1361 GGGAGTACTGCAACCTGACGCAATGCTCAGAGC 1394
      |||||
Db 1444 ACGACTACTGTGATGTCCCTCAGTGGGCGCC 1477
      |||||
```

RESULT 11

```
US-09-206-059-29
; Sequence 29, Application US/09206059
; Patent No. 6201104
; GENERAL INFORMATION:
; APPLICANT: Macdonald, Nicholas
; APPLICANT: Sim, Kim Lee
; TITLE OF INVENTION: Angiogenesis-Inhibiting Protein Binding Peptides and
; TITLE OF INVENTION: Proteins and Methods of Use
```

```
FILE REFERENCE: 05213-0370
; CURRENT APPLICATION NUMBER: US/09/206,059
; CURRENT FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 29
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-206-059-29
```

```
Query Match 4.3%; Score 216.2; DB 4; Length 1134;
Best Local Similarity 80.1%; Pred. No. 7.5e-64;
Matches 254; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
```

```
Qy 92 CAGCAGACCTGAGCAAGCAAGCATGTGTCAGAGATTGCTACCATGTGATGACAGAGTT 151
      |||||
Db 818 CAGCAGACCTGAGCTAAACCTGTGTGTCAGAGACTGCTACCATGTGATGACAGAGCT 877
      |||||
Qy 152 ATCGAGGACGCTACGACACGCTGTGCACAGAGAGACTGCGCAAGCTTGGTCATCTATGA 211
      |||||
Db 878 ACCGAGGCAATCTCCACACACACAGAGAAAGATGTCATCTTGTGATCATATGA 937
      |||||
Qy 212 CACCAATCAATATATATGACACAGAGAAACTACCCAAATGCTGCTGTGATCATGA 271
      |||||
Db 938 CACCAATCAATATATATGACACAGAGAAAGATGTCATCTTGTGATCATGA 997
      |||||
Qy 272 ACTGAGGAATCAAGATGCTGTGAGCTCTTATGTTATACGAGGAGATCCGTTGTCA 331
      |||||
Db 998 ACTGAGGAATCAAGATGCTGTGAGCTCTTATGTTATACGAGGAGATCCGTTGTCA 1057
      |||||
Qy 332 GGTGGAGGATCTGCAACCTGAGCAATGCTCAGAGAGAGGAGGAGGAGGAGGAGGAG 391
      |||||
Db 1058 GGTGGAGGATCTGCAACCTGAGCAATGCTCAGAGAGAGGAGGAGGAGGAGGAGGAG 1117
      |||||
Qy 392 CGACTGTACCCGGTT 408
      |||||
Db 1118 CGCTGTGTGCTGCTT 1134
      |||||
```

RESULT 12

```
US-08-441-370-1
; Sequence 1, Application US/08441370
; Patent No. 5721138
; GENERAL INFORMATION:
; APPLICANT: Lamn, Richard M.
; TITLE OF INVENTION: Apolipoprotein(A) Promoter and
; TITLE OF INVENTION: Regulatory Sequence Constructs and Methods of Use
; Patent No. 5721138
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIDUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/441,370
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/991,849
; FILING DATE: 15-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7627-003
```

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 790-9090
 TELEFAX: 212 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2542 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-08-441-370-1

Query Match 1.8%; Score 92.4; DB 1; Length 2542;
 Best Local Similarity 98.9%; Pred. No. 2, 2e-21;
 Matches 93; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTGGGATTGGGACACACTTTCGACACTGCTGCCAGTCCCAAAATGGAATAGGAA 60
 Db 1398 CTGGGATTGGGACACACTTTCGACACTGCTGCCAGTCCCAAAATGGAATAGGAA 1457
 Qy 61 GTGGTCTTCTACTCTTTTATTTCTGAATCAG 94
 Db 1458 GTGGTCTTCTACTCTTTTATTTCTGAATCAG 1491

RESULT 13
 US-08-985-526-22

; Sequence 22, Application US/08985526
 ; Patent No. 6080728

GENERAL INFORMATION:
 APPLICANT: Mixson, James A
 TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA
 TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
 TITLE OF INVENTION: THERAPY
 NUMBER OF SEQUENCES: 43
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Connolly, Bove, Lodge, & Hutz
 STREET: 1220 Market Street, P.O. Box 2207
 CITY: Wilmington
 STATE: Delaware
 COUNTRY: U.S.A.
 ZIP: 19899
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/985,526
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/608,845
 FILING DATE: 16-JUL-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: McMorrow Jr., Robert G
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (302) 658-9141
 TELEFAX: (302) 658-5613
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 645 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-985-526-22

Query Match 1.4%; Score 67.6; DB 3; Length 645;
 Best Local Similarity 55.0%; Pred. No. 3e-13;
 Matches 133; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 464 AGGAGTCTACATGTGTAATGACAGACTTATGAGGACACATCTCCACCACCTGTACAG 523

Db 266 AGGATCATGATCTCTGACTGAGAAAGATGAGGCAAAATCTCCAGACCATCTGTG 325
 Qy 524 GAGAACCTCCGACAGCTTGTCATATGACACACACTGCGATAGTGGACCCAGAAAT 583
 Db 326 GACTTGACTCCGAGCTTGTCATATGACACACACTGCGATAGTGGACCCAGAAAT 385
 Qy 584 ACTACCCAAATGCTGCTTGCATGATCACTGACAGAAATCCAGATGCTGCGAGCTC 643
 Db 386 AATTCACAGCAGAACTGAGATGATATGTCACACACCTGACGAGGAGCCAGAGC 445
 Qy 644 CTATGTATGATGAGGATCCGCTGTCAGTGGAGTATGATGACATGAGCAATGCT 703
 Db 446 CCTGCTCTTCACAAAGACCCACCAACGCTGGAAATCTGTGACATCCCCGCTGCA 505
 Qy 704 CA 705
 Db 506 CA 507

RESULT 14
 US-08-985-526-24

; Sequence 24, Application US/08985526
 ; Patent No. 6080728

GENERAL INFORMATION:
 APPLICANT: Mixson, James A
 TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA
 TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
 TITLE OF INVENTION: THERAPY
 NUMBER OF SEQUENCES: 43
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Connolly, Bove, Lodge, & Hutz
 STREET: 1220 Market Street, P.O. Box 2207
 CITY: Wilmington
 STATE: Delaware
 COUNTRY: U.S.A.
 ZIP: 19899
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/985,526
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/608,845
 FILING DATE: 16-JUL-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: McMorrow Jr., Robert G
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (302) 658-9141
 TELEFAX: (302) 658-5613
 INFORMATION FOR SEQ ID NO: 24:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1284 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-985-526-24

Query Match 1.4%; Score 67.6; DB 3; Length 1284;
 Best Local Similarity 55.0%; Pred. No. 4.4e-13;
 Matches 133; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 806 AGGAGTCTACATGTGTAATGACAGACTTATGAGGACACATCTCCACCACCTGTACAG 865
 Db 905 AGGATCATGATCTCTGACTGAGAAAGATGAGGCAAAATCTCCAGACCATCTGTG 964
 Qy 866 GAGAACCTCCGACAGCTTGTCATATGACACACACTGCGATAGTGGACCCAGAAAT 925
 Db 965 GACTTGACTCCGAGCTTGTCATATGACACACACTGCGATAGTGGACCCAGAAAT 1024

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 06:40:43 ; Search time 758 Seconds
(without alignments)

4405.804 Million cell updates/sec

Title: US-09-923-515-3_COPY_1_5000

Perfect score: 5000
1 cctggatctggagacacattt.....caagcttgctcatatgac 5000

Sequence:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 478924 seqs, 33395956 residues

Total number of hits satisfying chosen parameters: 957848

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Published_Applications_NA:*
2: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
3: /cgn2_6/ptodata/2/pubpna/PC1_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/PC1US_PUBCOMB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1855.6	37.1	10422	9 US-09-870-759-125	Sequence 125, Appl
2	236.6	4.7	2497	10 US-09-946-893-1	Sequence 1, Appl
3	216.2	4.3	1134	10 US-09-873-676-29	Sequence 29, Appl
4	165.4	3.3	431	10 US-09-864-761-13941	Sequence 13941, A
5	145.2	2.9	422	10 US-09-833-381-252	Sequence 252, App
6	125	2.5	168	10 US-09-864-761-30505	Sequence 30505, A
7	118.4	2.4	160	10 US-09-864-761-20992	Sequence 20992, A
8	105.6	2.1	408	10 US-09-960-352-1535	Sequence 1535, Ap
9	93.2	1.9	1164	10 US-09-880-107-2422	Sequence 2422, Ap
10	90.2	1.8	242	10 US-09-960-352-13971	Sequence 12971, A
11	82.8	1.7	362	10 US-09-864-761-4339	Sequence 4239, Ap
12	67.6	1.4	645	12 US-10-036-869-22	Sequence 22, Appl
13	67.6	1.4	1284	12 US-10-036-869-24	Sequence 24, Appl
14	62.8	1.3	1344	9 US-10-081-309-1	Sequence 1, Appl
15	62.8	1.3	5898	10 US-09-880-107-3708	Sequence 3708, Ap
16	61.8	1.2	780	9 US-09-870-759-96	Sequence 96, Appl
17	58.8	1.2	402	10 US-09-960-352-13108	Sequence 13108, A
18	58	1.2	245	10 US-09-960-352-4911	Sequence 4911, A
19	57.6	1.2	214	10 US-09-960-352-1590	Sequence 1590, Ap

20	39.8	0.8	383	10 US-09-960-352-12502	Sequence 12502, A
21	38.2	0.8	176	10 US-09-864-761-31962	Sequence 31962, A
22	38.2	0.8	453	10 US-09-864-761-15448	Sequence 15448, A
23	36.8	0.7	216	10 US-09-960-352-12269	Sequence 12269, A
24	35.6	0.7	9980	10 US-09-880-107-3352	Sequence 3352, Ap
25	34.4	0.7	417	10 US-09-960-352-8663	Sequence 8663, Ap
26	33.6	0.7	461	10 US-09-964-824A-112	Sequence 112, Ap
27	33.6	0.7	10422	9 US-09-870-759-125	Sequence 125, App
28	33.2	0.7	11990	10 US-09-969-708-559	Sequence 569, App
29	33	0.7	440	10 US-09-895-828-293	Sequence 293, App
30	32.8	0.7	529	10 US-09-998-598-2598	Sequence 2598, Ap
31	32.8	0.7	556	10 US-09-919-580-316	Sequence 316, App
32	32.8	0.7	561	10 US-09-919-580-556	Sequence 556, App
33	32.8	0.7	564	10 US-09-919-580-310	Sequence 310, App
34	32.8	0.7	752	10 US-09-919-580-859	Sequence 859, App
35	32.8	0.7	756	10 US-09-919-580-821	Sequence 821, App
36	32.8	0.7	873	10 US-09-880-107-2349	Sequence 2349, Ap
37	32.8	0.7	1000	10 US-09-925-301-180	Sequence 180, App
38	32.6	0.7	372	10 US-09-960-352-7521	Sequence 7521, Ap
39	32.6	0.7	401	10 US-09-864-761-14212	Sequence 14212, A
40	32.6	0.7	8034	10 US-09-764-860-890	Sequence 890, App
41	32.4	0.6	418	10 US-09-983-965-713	Sequence 713, App
42	32.2	0.6	234	10 US-09-960-352-5889	Sequence 5889, Ap
43	32.2	0.6	389	10 US-09-960-352-6628	Sequence 6628, Ap
44	32.2	0.6	411	10 US-09-960-352-9849	Sequence 9849, Ap
45	32	0.6	2000	9 US-09-938-842A-2905	Sequence 2905, Ap

ALIGNMENTS

RESULT 1
US-09-870-759-125
: Sequence 125, Application us/09870759
: Patent No. US20020177551A1
: GENERAL INFORMATION:
: APPLICANT: TERMAN, David S
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
: FILE REFERENCE: 870759
: CURRENT APPLICATION NUMBER: US/09/870,759
: CURRENT FILING DATE: 2002-01-14
: PRIOR APPLICATION NUMBER: US 60/208,128
: PRIOR FILING DATE: 2000-05-30
: NUMBER OF SEQ ID NOS: 166
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 125
: LENGTH: 10422
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (6667)..(10176)
: OTHER INFORMATION:
US-09-870-759-125

Query Match	37.1%	Score 1855.6;	DB 9;	Length 10422;
Best Local Similarity	83.0%	Pred. No. 0;		
Matches 2138;	Conservative	0;	Mismatches 414;	Indels 24;
			Gaps	1;
OY 1623	TGGCTGTATCTGAACTACCTCAGAACTCAGATGCTGGAGCTCCCTATTGTTATAC	1682		
DB 6657	TGGCTGTATCTGAACTACCTCAGATGCTGGAGCTCCCTATTGTTATAC	6716		
OY 1663	GAGGATCCCGGTGTCAAGTGGAGTACTGCAACTGACCAATGCTCAGACGACAGAG	1742		
DB 6717	GAGGATCCCGGTGTCAAGTGGAGTACTGCAACTGACCAATGCTCAGACGACAGAG	6776		
OY 1743	GACTCCGCTGCGCTCGCTCGACTGTTACCCGGTTCGAAGCTTGAGAGCTCCCTCGAACA	1802		
DB 6777	GACTCCGCTGCGCTCGCTCGACTGTTACCCGGTTCGAAGCTTGAGAGCTCCCTCGAACA	6836		
OY 1803	AGCACCGACTGAGCAAAAGGCTGGAGTGCAGAGATGCTTACCATGTAATGACAGATTA	1862		

Db 6837 AGCAACAACGTGACAAAGGCGCTGGGGTGCAGAGTCTTACCACGGAATGACAGATTG 6896
QY 1863 TCGAGGACATACTCCACACCTGTACAGAGAAGAACTGGCAAGCTTGGTATCTATGAC 1922
Db 6897 TCAAGGCACTACTCTATCTACTGTACACAGAGAAGAACTGGCAAGCTTGGTATCTATGAC 6956
QY 1993 ACCACATCCGATGTGCGGAGCCCGAGAAATCTACCCAAATGCTGGCTGTATCTATGAACTA 1982
Db 6957 ACCACATCCGATGTGCGGAGCCCGAGAAATCTACCCAAATGCTGGCTGTATCTATGAACTA 7016
QY 1983 CTGCGAGAAATCCAGATCTGTGAGAGCTCTTATTTGTATACGAGGATCCCGGTGTACG 2042
Db 7017 CTGCGAAATCCAGATCTGTGAGAGCCCTTGGTGTATACAAAGATCCCGGTGTACG 7076
QY 2043 GTGGGATGACTGCACACTGTACGCAATGCTACAGCGAAGAGGATCCGCTGGCGCTCC 2102
Db 7077 GTGGGATGACTGCACACTGTACGCAATGCTACAGAGAAATGAGATGCTCTCTCTCTCC 7136
QY 2103 GACTGTTACCCGGGTTCAGAGCCCTAGAGGCTCTCTTCGGAACAGCAACCGACTGAGCAAG 2162
Db 7137 GAATGTTATCTGGCTCCAGAGCTTAGAGGCTTTTGTGAACAGCACTGACTAGGAAAC 7196
QY 2163 GCCTGGGCTGAGAGAGTGTACCATGGTAAATGACAGAGATTATCGAGGACATATCTCAC 2222
Db 7197 CCGCGGGTACAGAGACTGTACTACTATTAAGACAGAGTTACCGAGGACATATCTCAC 7256
QY 2223 CACTGTACAGAGAAACCTGGCAAGCTTGTATCATGTATGACACACACCTGCATAGTGC 2282
Db 7257 CACTGTACAGAGAAACCTGGCAAGCTTGTATCATGTATGACACACACAGATAGTGC 7316
QY 2283 GACCCAGAAATACTACCCAAATGCTGGCTTATGATGAACTACTGCGAGAAATCCAGATGC 2342
Db 7317 GACCCAGAAATACTACCCAAATGCTGGCTGACAGAGAACTACTGCGAGAAATCCAGATGC 7376
QY 2343 TGTGGAGCTCTCTTATGTATGAGAGGATCCCGGTGACAGGTGGGAGTACTGCAACT 2402
Db 7377 TGTGATTCGGCTTGGTGTACACATGATGATCCAGTGTGAGGTGGGAGTACTGCAACT 7436
QY 2403 GACGCAATGCTCAGACGACAGAGGAGTCCGCTGGCGCTCCGACTGTAAACCCGGTTC 2462
Db 7437 GACGCAATGCTCAGTGTGACAAATCAAGTCTCTGCAACTCTCAGGTGTCCCAATGC 7496
QY 2463 AAGCTTAGAGGCTCTTCGGAACAGACGAGTGTAGCAAAAGGCGTGGGTGAGAGTGC 2522
Db 7497 AAGCAGAGAGGCTTCTTCTGAGAGAGACCAAGGAGCAAAAGCCCGGGGTCCAGGATTG 7556
QY 2523 CTACCATGTGTAATGACAGAGATTATGAGGACATACTCCACACCTGTACAGAGAAAG 2582
Db 7557 CTACCATGTGTAATGACAGAGATTATGAGGCTCATTTCTACACACTGTACAGAGAAAG 7616
QY 2583 CTGCGAAGCTTGTGATCTATGACACACACTGTGATAGTGGAGCCCGAGAAATACTACCC 2642
Db 7617 ATGTGATCTTGGTCTCTATGACACCACTGGCATCAGAGAAACCAAGAAATATATATCC 7676
QY 2643 AAATGTGCTGTATGATCAATGAACTACTGCAAGAAATCAGAAATGCTGGAGCTCTTATG 2702
Db 7677 AAATGTGCTGTATGATCAATGAACTACTGCAAGAAATCAGAAATGCTGGAGTACTGCTG 7736
QY 2703 TTATACAGAGGATCCCGGTCTCAGGTGGAGTACTGCAAACTGAGCGCAATGCTCAGAGCC 2762
Db 7737 TTATACAGGATCCCAATCTCAGATGAGTACTGCAACTGACACATGCTGCAGTGCAG 7796
QY 2763 AGAAGGACTGCGGTGCGGCTCCGACTGTACCCCGGTTCCAAAGCTTAGAGGCTCTCTTC 2822
Db 7797 AGAATCAAGTGTCTTGTGCAAGCTCAC-----GGCTGTCTTC 7832
QY 2823 CGAACAAGACCGACTGAGCAAAAGCGCTGGGTGACAGAGTGTCTACCATGTTATGAGACA 2882
Db 7833 TGAACAAGACCGCAAGAGAAAGCCCAAGTCCAGAGCTGTACCATGTTATGAGAGACA 7892
QY 2883 GAGTTATCGAGGACATATCTCCACACTGTGACAGAGAAAGAACTGCGCAAGCTTGGTATC 2942
Db 7893 GAGTTATCGAGGCTCATTTCTCCACACTGTATACAGAGAAAGACATGCACTGTGGTCTTC 7952

QY 2943 TATGACACCAACTGTCATAGTCTGACCCAGAAATACTACCCAAATGCTGGTATCAT 3002
Db 7953 TATGACACCAACTGTCATAGAGAACACAGAAATACTACCCAAATGCTGGTATCAT 8012
QY 3003 GAACTACTGCAAGAAATCCAGATGCTGTGGAGCTCTTATGTTATACAGAGGATCCCGG 3062
Db 8013 GAACTACTGCAAGAAATCCAGATGCTGTGGAGCTCTTATGTTATACAGAGGATCCCGG 8072
QY 3063 TGTGAGTGGAGTACTGCAACTGACGCAATGCTCAGAGCAGAGAAAGGACTGCGCTGC 3122
Db 8073 TGTGAGTGGAGTACTGCAACTGACGCAATGCTCAGAGTATGATGAAATCAACTCTCTCCAC 8132
QY 3123 GCCTCCGACTGTACCCCGGTTCCAAAGCTTAGAGGCTCTTCCGAAACACACCGACTCA 3182
Db 8133 AACTCCACAGGTTGCCAGAGTCCAGAGCAGAGCTTCTCTTGTGAAGAACACCACTGCA 8192
QY 3183 GCAAAGGCGCTGGGGTGCAGAGTGTACCATGTTATGAGCAGAGATTATGAGGACATTA 3242
Db 8193 AAGCAGCACTGGGGTCCAGAGCTGTACCGAGTGTATGACAGAGTTATGAGGACACT 8252
QY 3243 CTCCACCACTATCACAGAGAAACATGTCACTGTGGCTGTATGACACCACTATGCGCA 8312
Db 3303 TAGTGGAGCCCAAGATCTACCCAAATGCTGGCTGTATGATGAACTAGTCAGAGAAATCC 3362
QY 8313 TCGAGAGATCCCATTAATCTATCCAAATGCTGGCTGTACCAAGAACTACTGCGAGAAATCC 8372
QY 3363 AGATGCTGTGCAAGCTCTTATTTATACAGAGGATCCCGGTGTACAGTGGGAGTACTG 3422
Db 8373 AGATGCTGAGATTTGCGCTTGTGTATACCATGATGATCCAGTGTGAGTGGAGTACTG 8432
QY 3423 CAACCTGACGAAATGTCTCAGAGCAGAGAGGAGTCCGCTGGCGCTCCGACTGTACCCC 3482
Db 8433 CAACCTGACGAAATGTCTCAGAGCAGAGTGTCTCAGAACTCTCCACACTGCGCCC 8492
QY 3483 GGTTCAGAGCTTAGAGGCTCTTCGGAACAGCACCGCACTGAGCAAAAGGCTGGGTGCA 3542
Db 8493 GGTTCAGAGCAGAGGCTCTCTCTGAAACAGCACCGCACTGAGCAAAAGGCTGGGTGCA 8552
QY 3543 GAGTGTACCAAGTGTATGAGACAGAGTATTCAGAGGACATCTCCACACCTGTCCAGG 3602
Db 8553 GGATGTCTACCATGTGTGAGAGGAGTATTCAGAGGATATCTCCACCACTGTCCAGG 8612
QY 3603 AAGAACTGCGCAAGCTTGTCTATGACACCACTGCACTAGTGTGAGAGCCCAAGATA 3662
Db 8613 AAGGACCTGTCAATCTTGTGTATGATATACCACTGCACTAGGACATCAGAGGCCCAAGAAA 8672
QY 3663 CTACCCAAATGCTGGCTTGTATGAACTACTGAGAGAAATCCAGATGCTGTGCACTGC 3722
Db 8673 CTACCCAAATGCTGGCTGTGAGAGGAACTACTGAGAGAAATCCAGATGCTGTGCAAAAC 8732
QY 3723 TTATGTTATGAGAGGATCCCGGTGTACAGTGGAGTACTGCAACTGCAAGCAATGCTC 3782
Db 8733 CTGGGTGTACCAACCATCCGTGTGTGAGTGGAGTACTGCAATGCTGACAAATGCTC 8792
QY 3783 AGACGAGAAAGGAGTCCGCTGCGCTCCGACTGTAAACCCGGTTCAGAGCTTAGAGGC 3842
Db 8793 AGAAACAGAAATCAGAGTGTCTAGAGACTCCACAGTGTGTTCCAGTTCCAAAGCATAGAG 8852
QY 3843 TCTTTCGGAACCAACGAGGCTGTGAGCAAAAGGCGTGGGTGTCAGAGAGTGTACCATGTTA 3902
Db 8853 TCAATTTGAGACACACCAACTGAGCAAAACCCGTGTGTGTCGCAAGTGTACCATGTTA 8912
QY 3903 TGGACAGATTATGAGGACATCTCCACACTGTGTACAGAGAAAGCACTGCAAGCTTG 3962
Db 8913 TGGCAGAGATTATGAGGACATCTCCACACTGTGTACAGAGAAAGCAATGCAATCTTG 8972
QY 3963 GTCACTTATGACACCACTGCACTAGTGTGGAGCCCAAGAAATACTACCCAAATGCTGTG 4022
Db 8973 GTCACTTATGACACCACTGCACTAGTGTGGAGCCCAAGAAATACTACCCAAATGATGGCT 9032

QY	4023	GATATTAATCTACGACGAAATCCATATCTGTGGACGCTCTTTATTTTATATACAGGGA	4082
Db	9033	GACATTAATCTACTGCAGAAATCCAAATATCCATACAGGCCCTTGTGTTTATCATATGA	90922
QY	4083	TCCCGGTGTAGGTGGGAGATCTGCAACCTGACAGCAATCTTATAGACAGAAAGGACATGC	41424
Db	9093	CCCCAGATATAGGTGGGAGATCTGCAACCTGACAGCATCTTATAGACAGAAAGGACATGT	91522
QY	4143	CGTGCAGCTCCGACTGTTATCCCCGGTTCACAACTTAGAGGCTCTTCCGAACAG	4198
Db	9153	GGTGTCTCTCCGACTGTATATCCAGGTTCACAACTTAGAGGCTCTTCTTGAACAG	9208

RESULT 2

US-09-946-893-1
Sequence 1, Application US/09946893
Patent No. US20020072494A1
GENERAL INFORMATION:
APPLICANT: Cao, Yihai
TITLE OF INVENTION: Materials and methods relating to endothelial cell growth
TITLE OF INVENTION: Inhibitors
FILE REFERENCE: Newburn
CURRENT APPLICATION NUMBER: US/09/946, 893
CURRENT FILING DATE: 2001-09-05
PRIOR APPLICATION NUMBER: US 60/230,893
PRIOR FILING DATE: 2000-09-05
NUMBER OF SEQ ID NOS: 11
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 2497
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (50)..(2482)
US-09-946-893-1

Query Match	4.78;	Score 236.6;	DB 10;	Length 2497;
Best Local Similarity	58.68;	Pred. No. 3.3e-67;		
Matches 547; Conservative	0;	Mismatches 315;	Indels 72;	Gaps 5

QY	809	AGTCTACCATGGTATATGACACAGATTATCGAGGCACATATCCACACTGTCCACAGAA	868
Db	870	AGTCTCGAAGGGGAACAGGGTAAACTATCGCGGAATGTGGCTGTAACTGTGCCGGGC	929
QY	869	GAACCTGCCAAGCTTGCTCATCTATGACACACACACTCGCATGTGTGGACCCCAATATCT	928
Db	930	ACACTGTGCACACTGGATGAGTGCACACACCCCTTACACATATAACAGACACCCAGAAACT	989
QY	929	ACCCAAATGCTGCTTGATCATGACTACTGACAGATTCAGATGCTGTGGCAGCTCCT	988
Db	990	TCCCTGCACAAAATTTGATGAATAACTACTGCCGCAATCTGACGGAAAMAGGCCCAT	1049
QY	989	ATTGTATAGAGGAGATCCCGGTGTAGGTGGGAGATATGCAACCGACGCATATCTCAG	1048
Db	1050	GGTCCATPACACCAACAGCCAAAGTGGGGGAGTACTGTAAATACCTG-----	1100
QY	1049	ACGCAGAGGAGACTGCCGCTCCGACTTTTACCCTGGTTCCAAAGCTAGAGGCTC	1108
Db	1101	-----CCTGTGACTCTCTCCACGATATCCAGS---GAACAAT	1133
QY	1109	CTTCCGAAACAGCACCCGACTGACACAAAGGCTGTGGGTGAGAGATCTTACCATGTATG	1166
Db	1134	TGCGTCCACAGCACACCTGACTTAAACCCTGTGTGTCCAGACATGCTCAACATGGTGTATG	1193
QY	1169	GACAGATTTATCGAGGACATATCTCCACACTGTGTCCAGAGAAAGAACTCCCAAGCTTGT	1228
Db	1194	GACAGAGCTACCGAGGACATCTCTCCACACACACACAGAAAGATGTACGTCTTGGT	1255
QY	1229	CATCTATGACACACACTCTGCATATGTGGACCCAGATATCTACCAATATGCTGCCTTGA	1288
Db	1254	CATCTATGACACACACCTCCGACACAGAGAGCCCAATAAATATACCAATATGCTGTGCCTGA	1313

[illegible]

RESULT 3

```

US-09-873-676-29
? Sequence 29, Application US/09873676
? Patent No. US20020077289A1
? GENERAL INFORMATION:
? APPLICANT: Macdonald, Nicholas J.
? APPLICANT: Sim, Kim L.
? TITLE OF INVENTION: Angiotensin and Endostatin Binding Proteins and Methods of Use
? FILE REFERENCE: 05213-0378 (43170-25333)
? CURRENT APPLICATION NUMBER: US/09/873,676
? CURRENT FILING DATE: 2001-06-04
? PRIOR APPLICATION NUMBER: US 60/209,065
? PRIOR FILING DATE: 2000-06-02
? PRIOR APPLICATION NUMBER: US 60/289,387
? PRIOR FILING DATE: 2001-05-08
? NUMBER OF SEQ ID NOS: 123
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 29
? LENGTH: 1134
? TYPE: DNA
? ORGANISM: Homo sapiens
US-09-873-676-29

```

Query Match	4.38;	Score 216.2;	DB 10;	Length 1134;
Best local Similarity	80.18;	Pred. No. 1e-60;		
Matches 254; Conservative	0;	Mismatches 63;	Indels 0;	Gaps 0;

QY	92	CAGCAGCACC	TGAGCAAA	GGCATGTG	GTCCAGGATT	GCTACCAT	TGCTGAGCAGACTT	153
Db	818	CAGCAGCACC	TGAGCTAA	CCCCCTG	TGGTCG	CAGGACTG	CTACCATGTGATGAGCAGAGCT	877
QY	152	ATGAGGCA	AGTACTCC	ACCATGTG	CAAGAGAG	AGCTGCC	AGCTTGTGTCATCTATGA	211
Db	878	ACCGAGCA	ATCTCTCC	ACACACAC	ACACAGGAA	GAAGTGCAT	GTGTGGTCATCTATGA	937
QY	212	CACGACAT	CAACTAA	TAGGAC	CAAGAAA	CTTCCCAAT	TGCTGATGATCAACT	274
Db	938	CACCAAC	ACGGGCA	CCGAGAG	AGCCCCG	AGAAACTT	ACCMAATGTGCTCTGCACATGAACCT	997
QY	272	ACTGCAAG	AATCCAG	ATGCTGTG	GGCAGCTCT	CTATTGTAT	TACGAGGAATCCCGGTGTCA	331

Accession	Sequence	Length
Db	ACTGCAGGAATCCGAGATGCCGATTAAGGCCCCCTGGTGTTCACACAGACCCAGACTC	1057
Qy	GGTGGGAGTACTGCACACCTGACCAATGCTCAGACAGCAGAGGAGATGCGCTGCGGCTC	391
Db	GGTGGGAGTACTGCACCACTGAAAAAATGCTCAGAGACAGACGAGATGTGTAGACACTC	1117
Qy	CGACTGTTAACCCCGTT	408
Db	CGCTGTGTGTCCTCTT	1134

```

RESULT 4
US-09-864-761-13941
Sequence 13941, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wenshang
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomica-x-1
CURRENT FILING DATE: 2001-05-23
PRIORITY FILING DATE: 2001-02-04
PRIORITY FILING DATE: 2000-02-04
PRIORITY FILING DATE: 2000-05-26
PRIORITY FILING DATE: 2000-08-03
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 13941
LENGTH: 431
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO ALI09933..8
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
US-09-864-761-13941

```

[illegible]

RESULT 5
 US-09-833-381-252
 : Sequence 252, Application US/09833381
 : Patent No. US20020132090A1
 : GENERAL INFORMATION:
 : APPLICANT: Robison, Keith E.
 : TITLE OF INVENTION: No US0200132090A1el Nucleic Acid and Protein Homologs
 : FILE REFERENCE: 5800-119
 : CURRENT APPLICATION NUMBER: US/09/833,381
 : CURRENT FILING DATE: 2001-04-11
 : PRIOR APPLICATION NUMBER: 09/516,448
 : PRIOR FILING DATE: 2000-02-29
 : NUMBER OF SEQ ID NOS: 2050
 : SOFTWARE: FASTSEQ for Windows Version 3.0
 : SEQ ID NO 252
 : LENGTH: 422
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-09-833-381-252

Query Match	2.98;	Score 145.2;	DB 10;	Length 422;
Best Local Similarity	87.48;	Pred. No. 1.2e-37;		
Matches 159; Conservative	0;	Mismatches 23;	Indels 0;	Gaps 0;

QY	255	TGGCTTGATCATGACACTACTCTCAGAGATCCAGATAGGTGTGGAGAGCTCCCTATATGTTATAC	31.4
Db	150	TGACTTGATCATGAACACTACTCTCAGAGATCCAGATAGGTGTGGAGAGCTCCCTGTGTTATAC	20.9
QY	315	GAGGAGATCCCGGTGTACAGTGGAGAGTACTCTCAACTGACGCATGCTCAGACGAGAAG	37.4
Db	210	GATGATATCCCAAGTCAATATGGAGATACTCTCAACTGACGCATGCTCAGACGAGAAG	26.9
QY	375	GACTGCGCTGGCGGCTCCGACTGTTACCCGGTTCGAAGCTTAGAGGCTCCTTCCGAACA	43.4
Db	270	GACTGCAATCGTGCTCTGACTGTTATCCGGTTCGAAGCTTAGAGGATCCTTCCAAACC	32.9
QY	435	AG 436	
Db	330	AG 331	

```

RESULT 6
US-09-864-761-30505
; Sequence 30505, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEOTIC ACID PROBES USEFUL FOR

```

```

; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30505
; LENGTH: 168
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL109933.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: NT HIT: M86879.1, EVALUATE 2.00e-82
; OTHER INFORMATION: EST_HUMAN HIT: R07514.1, EVALUATE 2.00e-30
; OTHER INFORMATION: SWISSPROT HIT: P08519, EVALUATE 8.00e-22
; US-09-864-761-20992

Query Match      2.58; Score 125; DB 10; Length 168;
Best Local Similarity 100.0%; Pred. No. 3.1e-31;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 130 TACCACTGTGATGAGAGTATTCAGGACGATCCACCACTGTCCACAGGAGACC 189
    |||||||
DB 1 TACCACTGTGATGAGAGTATTCAGGACGATCCACCACTGTCCACAGGAGACC 60
OY 130 TGCACAGCTTGATCATCTATGACACGCATCAACATATATGACACAGAAAATCTACCA 249
    |||||||
DB 61 TGCACAGCTTGATCATCTATGACACGCATCAACATATATGACACAGAAAATCTACCA 120
OY 250 AATGC 254
    |||||
DB 121 AATGC 125

RESULT 7
```

```

; US-09-864-761-20992/c
; Sequence 20992, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20992
; LENGTH: 160
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL109933.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 6.5
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.8
; OTHER INFORMATION: EST_HUMAN HIT: T89505.1, EVALUATE 4.00e-85
; OTHER INFORMATION: SWISSPROT HIT: P00747, EVALUATE 2.00e-21
; OTHER INFORMATION: NT HIT: M74220.1, EVALUATE 3.00e-85
; US-09-864-761-20992

Query Match      2.4%; Score 118.4; DB 10; Length 160;
```


Best Local Similarity 83.8%; Pred. No. 4.4e-29;
Matches 134; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY	95	CACACCTCAGCAAAAGCCATCTGGCTCCAGATATGCTACCTGGTATGTAGACAGAGTTATC	150
Db	160	CACACCTCAGCTAACCCCTGTGTGCCAGACCTCTCAACATGGTATGTAGACAGACTACCC	1010
QY	155	GAGGCACGACTCTCCACCACTGTCTACAGAAAGGACTCTGCCAAGCTTGTATCATCTATGACAC	214
Db	100	GAGGCACATCTCTCCACCACTCCACCAACAGAGAAAGGTGATAGCTTGTGTATCTATGACAC	41
QY	215	CACATCAACATTAATGAGACCAAGAAAACCTACCCCAATGCG	254
Db	40	CACACCGGCATCCAGAAAGCCCCAGAAAATCTACCCCAATGCG	1

```

RESULT 8
US-09-960-352-1535
? Sequence 1535, Application US/09960352
? Patent No. US20020137139A1
? GENERAL INFORMATION:
? APPLICANT: Warren, Wesley C.
? APPLICANT: Tao, Nengbing
? APPLICANT: Byatt, John C.
? APPLICANT: Mathalagan, Nageshpan
? TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
? FILE REFERENCE: 16511.006/37-21(10298)C
? CURRENT APPLICATION NUMBER: US/09/960.352
? CURRENT FILING DATE: 2001-09-24
? NUMBER OF SEQ ID NOS: 15112
? SEQ ID NO 1535
? LENGTH: 408
? TYPE: DNA
? ORGANISM: Bos taurus
? OTHER INFORMATION: Clone ID: 07-LTB34-018-01-E1-B11
US-09-960-352-1535

```

Query Match	2.1%;	Score 105.6;	DB 10;	Length 408;
Best Local Similarity	56.6%;	Pred. No. 1.2e-24;		
Matches 241; Conservative	0;	Mismatches 149;	Indels 36;	Gaps 1;

OY	142	GGACAGATTATCGAGGCAGCTACTCCACCACTGTGACAGAAAGACCTGCCAAGCTTGG	207
Db	13	GGGAAAAAATAATGGAAGGAGCAGTGACCGGTACCAGAGTCGGGGCACACAATGTACAGCGCTGG	72
OY	202	TCACTATGACACCACTACATCATATATAGAACCCAGAAAACCTACCCAAATGTGGCGCTG	261
Db	73	AGGAGACACACACCCCAACAAACACMACAGAGACCCCAGAAAACCTTTCATGACAAAACTTG	133
OY	262	ATCATGACTACTGCAAGAAATCCAGANTGTTGGCAGCTCTTATTTATTAACGAGGAT	321
Db	133	GAGGAAAAATACTGTCTGTAACCTTAATGAGGAGAAAGCCCCATGGGTATATCAAACCAAC	197
OY	322	CCCGGTGAGGTGGGATACTGCAACCTGACCAACCAATGATGACAGCAGAAAGGAGCTGCC	381
Db	193	AGCGAAATGATGGTAGGAATCTGCAACAATCCCGTCTGCG-----	231
OY	382	GTCGCGCCTCCGACTGTACCCCGGTTCCAAGCTTAGAGGCTCTTCCGAACAGACACG	441
Db	232	-----GAGTCTCTCCGTTATATCCAGGAAACGATGAGATGTCCCAATACCA	276
OY	442	ACTGACCAAAAGGCTTGGGGGTCCAGAGAATGCTACCATGTGTAATGACAGAGTTATTCAGAGC	501
Db	277	CCCGAAACAAATCTCGTGTCCGCCCAAGACATGCTTACCATATGTTAAAGGGCAAGATTAATTCAGAGC	336
OY	502	ACATATCTCCACACCTGTGCACAGAAAGCAATGGCCAAAGCTTGGTCATCTATGACACCAACAC	561
Db	337	ACCTGCTCCACTACTATATCAGAGAAAGAAATGCCAGTCTGTGGTCATGCATGACACCAACAC	396
OY	562	TGCGAT	567
Db	397	AGGCAT	402

RESULT 9
 US-09-880-107-2422
 Sequence 2422, Application US/09880107
 Patent No. US20020142981A1
 GENERAL INFORMATION:
 APPLICANT: Horne, Darci T.
 APPLICANT: Hockley, Joseph G.
 APPLICANT: Scherff, Dwe
 APPLICANT: Gene Logic, Inc.
 TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
 FILE REFERENCE: 44921-5028-WO
 CURRENT APPLICATION NUMBER: US/09/880,107
 CURRENT FILING DATE: 2001-06-14
 PRIOR APPLICATION NUMBER: US 60/211,379
 PRIOR FILING DATE: 2000-06-14
 PRIOR APPLICATION NUMBER: US 60/237,054
 PRIOR FILING DATE: 2000-10-02
 NUMBER OF SEQ ID NOS: 3950
 SOFTWARE: patentIn Ver. 2.1
 SEQ ID NO 2422
 LENGTH: 1164
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: Genbank Accession No. US20020142981A1 M93143
 US-09-880-107-2422

	Query Match	Similarity	Score	DB	length
Best Local	92.5%	92.5%	93.2	10	1164
Matches	98	Conservative	0	Mismatches	8
				Indels	0
				Gaps	0
QY	1	CTGGGATGGGACACACTTTCGGACATCGTGGCCAGTCCCAAAATGGAATAGAA	60		
Db	23	CTGGGATGGGACACACTTTCGGACATCGTGGCCAGTCCCAAAATGGAATAGAA	82		
QY	61	GTGGTCTTCTTACTTCTTTATTTTGTGAATCAGACACACTGAGC	106		
Db	83	GTGGTCTTCTTACTTCTTTATTTTGTGAATCAGAGTCAAGAGAGC	128		

```

RESULT 10
US-09-960-352-12971
; Sequence 12971; Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Ningbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalaagan, Nagappan
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 12971
; LENGTH: 242
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 55-LIB34-052-Q1-E1-F4
US-09-960-352-12971

```

Query Match	1.8%	Score 90.2	DB 10	Length 242
Best Local Similarity	81.9%	Pred. No. 9.8e-20		
Matches 104	Conservative	0	Mismatches 23	Indels 0
				Gaps 0
QY	92	CACGACACCTGTGACCAAGCCATGTGTGTCGAGATTCTTACCATGTGTGATGCACAGATT	151	
Db	113	CAGTACCACCCGACCAAAATCTCTGTGCCCCGAGACTCTTACCATGTGTGTAACGGCGACAGATT	172	
QY	152	ATCGAGGACAGTATTCACCAACTGTGCACAGGAAGAGCTCCAAAGTTGTGATCATATGA	211	

Db 266 AGAATGATGATGAGTGGAGAAAGATGAGGCAAAATCTCCAAGACATGTCTG 325
Qy 524 GAAGACCTGCGCAAGCTGTGTCATATGACACACACCTCGATGTGGAGCCAGAT 583
Db 326 GACTTACTGCGCAGCTGGATCTTCAGAGCCACATCTCATGTGATACATCTGCA 385
Qy 584 ACTACCCAAATGCTGCTTGATCATGACACTGACAGAAATCAATGCTGTGCGAGCTC 643
Db 386 AATTTCAGAGAAAGACCTGAGATGATTAATGCGACCAACCTGACGGGAGCCAGGC 445
Qy 644 CTATTGTATACGAGGATCCCGGTGTGAGTGGAGTACCTGCAACCTGACGCAATGCT 703
Db 446 CCTGTGCTTCACACAGCCACCAACAGCTGGGAATACGTGACATCTCCCGCTGCA 505
Qy 704 CA 705
Db 506 CA 507

RESULT 13

US-10-036-869-24
; Sequence 24, Application US/10036869
; Patent No. US20020151516A1

GENERAL INFORMATION:

APPLICANT: Mixson, James A
TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA
ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
THERAPY

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: Connolly, Bove, Lodge, & Hutz
STREET: 1220 Market Street, P.O. Box 2207
CITY: Wilmington
STATE: Delaware
COUNTRY: U.S.A.
ZIP: 19899

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/036,869
FILING DATE: 29-NO. US20020151516A1-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/985,526
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/608,845
FILING DATE: 16-JUL-1996

ATTORNEY/AGENT INFORMATION:
NAME: McMorrow Jr., Robert G

TELECOMMUNICATION INFORMATION:
TELEPHONE: (302) 658-9141
TELEFAX: (302) 658-5613

INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1284 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-10-036-869-24
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:

Query Match 1.48; Score 67.6; DB 12; Length 1284;
Best Local Similarity 55.08; Pred. No. 6, 5e-12;
Matches 133; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 806 AGAGTGTCTACATGATGATGACAGAGTATGAGCAGACATCTCCACACTGTGTACAG 865
Db 905 AGAATGATGATGATGATGAGAGAAAGATGAGGCAAAATCTCCAAAGCAGTGTG 964
Qy 866 GAAGACCTGCGCAAGCTGTGTCATATGACACACACTGCTGATGTGGAGCCAGAT 925

Db 965 GACTTGACCTGCGCAGGCTGGGATTCCTCAGAGCCACATGCTCATGTGATCTCCGCA 1024
Qy 926 ACTACCCAAATGCTGCTTGATCATGAACTATCTGCAAGAAATCCAGATGCTGTGCGAGCTC 985
Db 1025 AATTTCAGAGAAAGACCTGAGATGATTAATGCGACCAACCTGACGGGAGCCAGGC 1084
Qy 986 CTATTGTATACGAGGATCCCGGTGTGAGTGGAGTACCTGCAACCTGACGCAATGCT 1045
Db 1085 CCGTGTGCTTCACACAGCCACCAACAGCTGGGAATACGTGACATCTCCCGCTGCA 1144
Qy 1046 CA 1047
Db 1145 CA 1146

RESULT 14

US-10-081-309-1

; Sequence 1, Application US/10081309

; Publication No. US20030012775A1

; GENERAL INFORMATION:

; APPLICANT: Hoffmann-La Roche Inc.

; TITLE OF INVENTION: PEG Conjugates of NK4

; FILE REFERENCE: 20859

; CURRENT APPLICATION NUMBER: US/10/081,309

; CURRENT FILING DATE: 2002-06-04

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 1344

; TYPE: DNA

; ORGANISM: homo sapiens

US-10-081-309-1

Query Match 1.38; Score 62.8; DB 9; Length 1344;
Best Local Similarity 53.78; Pred. No. 2, 5e-10;
Matches 130; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy 124 GATTGTACCATGATGATGAGACAGAGTATGAGAGCACTGACCTGTCACAGGA 183
Db 817 GATGATCATCCAAAGCTGACAGAGAGAGGCTACAGGGGCTGTCAATACATTGGATGSA 876
Qy 184 AGAGCTGCGCAAGCTTGCTCATCTATGACACACATCAATATATGACACAGAAAAC 243
Db 877 ATTCATGATGACGCTGGGATTCAGTATCTCAGAGCATACATGACTCCTGAAT 936
Qy 244 TACCAAAATGCTGCTGATCATGAACTACTGACGAATCCAGATGCTGTGACGCTCT 303
Db 937 TTCAGTGCAGAGACCTACAGAAAATTAATGCGCGAATCCAGATGGTGTGATCACC 996
Qy 304 TATTGTTATGAGGAGATCCGGGTGTCAAGTGGAGTACTGCAACCTGACGCAATGCTCA 363
Db 997 TGGTGTATTCACAGATCAACATCCGAGTTGGTACGTGCTCCAAATTCCAAACTGT 1056
Qy 364 GA 365
Db 1057 GA 1058

RESULT 15

US-09-880-107-3708

; Sequence 3708, Application US/09880107

; Patent No. US20020142981A1

; GENERAL INFORMATION:

; APPLICANT: Horne, Darcil T.

; APPLICANT: Vockley, Joseph G.

; APPLICANT: Scherif, Uwe

; APPLICANT: Gene Logic, Inc.

; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

; FILE REFERENCE: 44921-5028-NO

; CURRENT APPLICATION NUMBER: US/09/880,107

; CURRENT FILING DATE: 2001-06-14

; PRIOR APPLICATION NUMBER: US 60/211,379

```
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3708
; LENGTH: 5898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X16323
US-09-880-107-3708
```

```
Query Match 1.3%; Score 62.8; DB 10; Length 5898;
Best Local Similarity 53.7%; Pred. No. 5.7e-10;
Matches 130; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
```

```
QY 124 GATTGCTACCATGATGATGAGCAGAGTTATCGAGGACGCTACTCCACCACTGTACAGGA 183
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1044 GAATGCATCCAAAGSTCAAGAGAGAGGCTACAGGGGCACCTGCCAATACATTGGAAATGA 1103
QY 184 AGGACCTGCCAAGCTTGCTCATCTATGACACCAATCAACATATATAGACACAGAAAC 243
    - - - - - ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1104 ATTCCATGTCAGGCTGGGATCTCTCAGTATCTCACAAGCATGACATGACTCTGAAAT 1163
QY 244 TACCCAAATGCTGCTGATCATGAATCTACTGCGAGAAATCCAGATGCTGTGCAGCTCT 303
    - - - - - ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1164 TTCAGTGCAGAGACTACGAGAAATTTACTGCCGAAATCCAGATGGGTCTGAAATCACC 1223
QY 304 TATGTTTATACAGAGGATCCGGGTGTCAGGTGGGAGTACTGCAACCTGACGCAATGCTCA 363
    - - - - - ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1224 TGGTGTTTTACCACTGATCCAAACATCCGAGTGTGTTACTGCTCCCAATTCCAAACCTGT 1283
QY 364 GA 365
    ||
DB 1284 GA 1285
```

Search completed: March 6, 2003, 15:50:52
Job time : 826 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 05:15:44 ; Search time 5999.5 Seconds
(without alignments)
13497.364 Million cell updates/sec

Title: US-09-923-515-3_COPY_1_5000
Perfect score: 5000
Sequence: 1 ctggagattggacacactt.....caagctgtcatcatatgac 5000

Scoring table:
IDENTITY_NOC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	229	4.6	666	10	AV654588 AV654588
2	215	4.3	336	14	R07567 yeg7h07.r1
3	203.4	4.1	461	14	R07514 yeg7h07.s1
4	199.4	4.0	353	10	AV652896 AV652896
5	191	3.8	304	14	H63293 yr49e03.r1
6	190.6	3.8	444	14	R91117 yp93g09.r1

7	189.4	3.8	456	12	BF654744
8	188.6	3.8	454	14	H49223
9	183.4	3.7	529	13	B1343595
10	175.4	3.5	608	10	AV655576
11	173.8	3.5	529	9	A1303081
12	170	3.4	393	9	AA504868
13	167	3.3	597	17	AV650465
14	159	3.2	597	17	AQ878410
15	152.6	3.1	431	9	A1328320
16	152.6	3.1	475	14	W99028
17	152.2	3.0	320	9	AA343990
18	150.8	3.0	701	13	B1219735
19	145.2	2.9	419	9	A1025149
20	145.2	2.9	422	9	A1023507
21	140.6	2.8	419	14	T89505
22	134.2	2.7	420	14	H94827
23	129.8	2.6	558	17	AQ518926
24	128.4	2.6	577	14	T81776
25	128.4	2.6	550	17	AQ533551
26	125.2	2.5	324	10	AM085134
27	124.4	2.5	288	14	T77573
28	123.2	2.5	788	10	AV695776
29	123	2.5	499	14	H54245
30	121.8	2.4	592	17	AO582455
31	120	2.4	750	10	AV695796
32	114.8	2.3	440	17	AQ074543
33	112	2.2	620	9	AA237452
34	110	2.2	675	17	AG037548
35	109.8	2.2	375	10	AV690608
36	109	2.2	361	10	AV651158
37	108.8	2.2	275	9	AA572527
38	107.6	2.2	465	14	T89681
39	106.6	2.1	469	17	AQ009509
40	104	2.1	292	14	T82089
41	103.6	2.1	493	17	AQ292872
42	103.4	2.1	483	14	T77537
43	102.8	2.1	518	9	A1545087
44	102	2.0	423	9	A1051709
45	97.8	2.0	448	17	AQ220996

ALIGNMENTS

RESULT 1	AV654588	666 bp	mrna	linear	EST 15-JAN-2002
LOCUS	AV654588				
DEFINITION	AV654588				
ACCESSION	AV654588				
VERSION	AV654588.1	GI:9875602			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 666)				
AUTHORS	Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X., Xiao,H., Qu,D., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,D., Hu,W., Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X., Hu,G., Gu,J., Chen,Z., and Han,Z.				
TITLE	Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver				
JOURNAL MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)				
COMMENT	Contact: Zenguan Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzy@chgc.sh.cn This clone is available at CHGC in Shanghai.				

Location/Qualifiers
1. .336

(Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pTR7 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fátima Bonaldo."

um (info@image.llnl.gov) for further information

(Pharmacla), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pTR7 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Felima Bonaldo."

The washu-Merck EST Project

[illegible]

```

Insert size: 2912
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Trace length: 2912      Std Error: 0.00
Seq primer: M13mp1
High quality sequence stop: 1.
Location/Qualifiers

```

```

/organism="Homo sapiens"
/db_xref="GDB:377767"
/db_xref="taxon:9606"
/cb_image="IMAGE:208636"
/clone_lib="Soares fetal liver spleen 1NFLS"
/sex="male"
/dev_stage="30 week post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/node="Organ: Liver and Spleen; Vector: p773D (Pharmacia) with a modified polyliner; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer 3', 5' AACTCGAAGATTAAATTAAGATCTTTTATTTTATTTT 3', double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified p773 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Felina Bonaldo."

```

	Query Match	3.8%	Score 191;	DB 14;	Length 304;
	Best Local Similarity	75.4%;	Pred. No. 7.1e-48;		
	Matches 227;	Conservative 0;	Mismatches 74;	Indels 0;	Gaps 0;
QY	108	AAGCATGTGGTCCAGGATGCTTACCATGATGATGACAGAGATTATGAGGACGCTACTC	167		
Db	4	AACCCGTGTGTCAGAGACTGCTACCATGTGTGATGAGACAGAGCTACCGAGGACATCTCC	63		
QY	168	CACCACTGTCACAGGAGGAACTGCCAAGCTTGGTCATCTATGACACACATCAACATTA	227		
Db	64	CANACCACACACAGGTAGTAAGACTNNCACTTGGTCATCTATGACACACACACCCGACACA	123		
QY	228	TAGGACACAGAAATATACCCAAATGCGTGTGATATGAACTACGACGGAATCCACA	287		
Db	124	GAGAACCCACGAAATATACCGAATGATGAGCTGACANTGACTACGACGAGGNNTCACA	183		

QY	288	TGCGTGGCAGCTCCCTTAATGTAATACGAGGATCCCGGTACAGTGGAGGAGTACGCA	347
Db	184	TGCGATAAAGGCCCCCTGGTNTTTTACACACACCCCGCTCAGGTGGAGTACTGCA	243
QY	348	CCTTACGCANTGCTCAGACGACGAGGAGCTCCGCTCCGCTCCGACTCTTAAACCCGGT	407
Db	244	CCTGAAAAATGCTCAGAGCAGACGACGAGTGTGTACACACTTCGGCTGTGTCCGGT	303
QY	408	T	408
Db	304	T	304

RESULT 6	
R01117	
LOCUS	444 bp mRNA linear EST 25-AUG-1995
DEFINITION	Y99309.r1 Soares fetal liver spleen INF15 Homo sapiens cDNA clone IMAGE:195040 5' similar to gb:X05199 PLASMINOGEN PRECURSOR (HUMAN)'; mRNA sequence.
ACCESSION	R01117
VERSION	R01117.1 GI:958657
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 444)
AUTHORS	Hillier,L., Clark,N., Dubuque,T., Ellison,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Ronfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wobidmann,P. and Wilson,R.
TITLE	The WashU-Merck EST Project
JOURNAL	Unpublished (1995)
COMMENT	Contact: Wilson RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1916
High quality sequence stops: 310
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1916 Std Error: 0.00
Seq primer: M13RPI
High quality sequence stop: 310.
Location/Qualifiers
1..444

```

/organism="Homo sapiens"
/db_xref="CDB:3764090"
/db_xref="taxon:9606"
/clone="IMAGE:195040"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACGGGAAGATTATTAAGAAGCTTTTTTTTTTTTTTTTTT 3'] double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Patima Bonaldo."
BASE COUNT      127 a      130 c      104 g      81 t      2 others
ORIGIN

```

Query Match 3.8%; Score 190.6; DB 14; Length 444;
Best Local Similarity 81.5%; Pzed.No. 1.2e-47;
Matches 220; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

FEATURES	Location/Qualifiers
SOURCE	1. 456
	/organism="Bos taurus"
	/db_xref="taxon:9913"
	/clone_1lb="MARC 3B0V"
	/tissue_type="pooled"
	/lab_host="DH10B"
	/note="Vector: pCMV SPORT6; site_1: NotI; site_2: SalI; library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendinosus muscle, and fetal longissimus muscle."
BASE COUNT	134 a 133 c 103 g 86 t
ORIGIN	
Query Match	3.8%; Score 189.4; DB 12; Length 456;
Best Local Similarity	81.2%; Pred.No. 2.8e-47;

```

FEATURES
source
location/Qualifiers
1.454
/organism="Homo sapiens"
/db_xref="CDB:3798132"
/db_xref="taxon:9606"
/clone_image="274351"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pRT3D (Pharmacia) with a modified polylinker; Site.1: Pac I; Site.2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTCGAGCAATTTATTTAAAGACTCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I

```


Insert length: 2323 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amer sham

```
FEATURES      High quality sequence stop: 385.
SOURCE        Location/Qualifiers
              1..393
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_lib="IMAGE:839703"
              /clone_1lb="Stratagene fetal retina 937202"
              /sex="mixed"
              /lab_host="SOLR (kanamycin resistant)"
              /note="Vector: pBluescript SK-, Site_1: EcoRI, Site_2:
              XhoI; Cloned unidirectionally. Primer: Oligo dT, pooled
              retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
              Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3',
              adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3'."

BASE COUNT    102 a 112 c 90 g 89 t
ORIGIN

Query Match    3.4%; Score 170; DB 9; Length 393;
Best Local Similarity 99.5%; Pred. No. 2.6e-41;
Matches 181; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 255 TGGCTTGATCATGAACTACTGAGAGATCCAGATGCTGTGCGAGCTCCTTATTTGTATAC 314
      |||||
DB 226 TGGCTTATCATGAACTACTGAGAGATCCAGATGCTGTGCGAGCTCCTTATTTGTATAC 167

QY 315 GAGGATCCCGGTGTGAGGTGGAGTACGACCACTGACCAAGCTGACGAGAGAG 374
      |||||
DB 166 GAGGATCCCGGTGTGAGGTGGAGTACGACCACTGACCAAGCTGACGAGAGAG 107

QY 375 GACTGCGGTGCGCGCTCCGACTGTATACCCGGTTCAGAGCCTAGAGGCTCCTTCGAGACA 434
      |||||
DB 106 GACTGCGGTGCGCGCTCCGACTGTATACCCGGTTCAGAG-CTAGAGGCTCCTTCGAGACA 48

QY 435 AG 436
      ||
DB 47 AG 46

RESULT 13
LOCUS      AV650465 687 bp mRNA linear EST 15-JAN-2002
DEFINITION AV650465 GIC Homo sapiens cDNA clone GICCFB02 3', mRNA sequence.
ACCESSION  AV650465
VERSION     AV650465.1 GI:9871479
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 687)
AUTHORS     Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
            Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,D., Hu,W.,
            Shen,K., Lu,G., Fu,G., Zhong,W., Xu,S., Gu,W., Huang,W., Zhao,X.,
            Hu,G., Gu,J., Chen,Z., and Han,Z.
            Insight into hepatocellular carcinogenesis at transcriptome level
            by comparing gene expression profiles of hepatocellular carcinoma
            with those of corresponding noncancerous liver
            Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
            21625106
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@hgc.sh.cn
            This clone is available at CHGC in Shanghai.

FEATURES
SOURCE      1..587
            Location/Qualifiers
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="GICCFB02"
            /clone_1lb="GIC"

/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT    185 a 174 c 181 g 138 t 9 others
ORIGIN

Query Match    3.3%; Score 167; DB 10; Length 687;
Best Local Similarity 77.8%; Pred. No. 3e-40;
Matches 214; Conservative 0; Mismatches 60; Indels 1; Gaps 1;

QY 163 TACTCCACCTACTGTTCACAGAGAGACCTGCCAAGTGTGTCATCTATGACACCCACATCAA 222
      |||||
DB 8 TCTCTCCACACACACACACAGAGAGAGTGTCACTCTGTGTCATCTATGACACCCACACCGG 67

QY 223 CATATATGAGACCAACAGAAACTACCCAAATGCTGCTTATCATGATCATGACAGAT 282
      |||||
DB 68 CACCAAGAGAGACCCAGAGAAACTACCCAAATGCTGCTTATCATGATCATGACAGAT 127

QY 283 CCAGATGCTGTGGAGCTCCTTATTTATGAGAGGATCCGGTGTGAGTGGAGTAC 342
      |||||
DB 128 CCAGATGCCAGATTAAGAGCCCTGCTGTTTACACAGACCCAGGCTGAGTGGAGTAC 187

QY 343 TGCAACCTG-ACGCAATGCTCAGACGAGAGGAGTGGCTGCGGCTCCGACTGTAC 401
      |||||
DB 188 TGCAACTGAGAAAAAATGCTCAGAGAACAGAGAGATGTTGTAGCACTCCGACTGTTGT 247

QY 402 CCCGGTCCAGGCTAGAGGCTCCTTCGAGCAG 436
      |||||
DB 248 CCGTCTCCAGATGTAGAGACTCTTCGAGCAG 282

RESULT 14
LOCUS      A0878410 597 bp DNA linear GSS 09-NOV-1999
DEFINITION HS_3069_A1.B10.T7C CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate-3069 Col-19 Row-C, DNA sequence.
ACCESSION  A0878410
VERSION     A0878410.1 GI:6309877
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 597)
            Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
            Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
            Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
            99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.hisc.washington.edu
            Plate: 3069 row: C column: 19
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 597.

FEATURES
SOURCE      1..597
            Location/Qualifiers
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="Plate-3069 Col-19 Row-C"
            /clone_1lb="CIT Approved Human Genomic Sperm Library D"
            /sex="male"
```

BASE COUNT	/note="Organ: sperm; Vector: pBelBAC11; BAC Clones in E-Coli DH10B"			
ORIGIN	154 a	134 c	132 g	166 t
				11 others

Query Match	3.28;	Score 159;	DB 17;	Length 597;
Best Local Similarity	98.18;	Pred. No. 8.2e-38;		
Matches 159; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

QY 435 ACCACCGAGTGAAGCAAGCCCTGGGGTGCAGAGTGTCTACCATGTGTAATGGACAGATT 494
|||||
Db 505 AGCACNGACTGACAAAAGGCGTTGGGGTGCAGAGTGTCTACCATGTGTAATGGACAGATT 446
|||||

495 TCGAGGACATCTCCACCACCTGTACAGGAAACCTGCCAAGCTTGGTCATCTATGAC 554
 |||||
 445 TCGAGGACATCTCCACCACCTGTACAGGAAACCTGCCAAGCTTGGTCATCTATGAC 386

0y 555 ACCACACTCGCATAGTCGGACCCAGAACTACCCCAATGC 596
 |||
 Db 385 ACCACACTCGCATAGTCGGACCCAGAACTACCTACCCCAATGC 344

RESULT 15	
AI529320	
LOCUS	
AI529320	431 bp
	mRNA
	linear
	FST 18-MAR-1994

DEFINITION
 uc022006.y1 sugarcow mouse liver lmda mus discuslms cdna clone
 IMAGE:18865966 5' similar to gb.X05199 PLASMINOGEN PRECURSOR (HUMAN
); gb.J04766 mouse plasminogen mRNA, complete cds (MOUSE);, mRNA
 sequence.

```

revision      11.2.2.20
version       AI529320.1  GI:443455
keywords      EST.
source        house mouse.

```

REFERENCE
1 (bases 1 to 431)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shiu, T., Jackson, Y., Cardenas, M., McCann, R.,

TITLE	The WashU-NCI Mouse EST Project 1999
JOURNAL	Unpublished (1999)
COMMENT	Contact: Marra M/WashU-NCI Mouse EST Project 1999

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

This clone is available royalty-free through LINL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:971290

FEATURES	Location/Qualifiers
source	1. .431
	/organism="Mus musculus"
	/taxonomy="ncrna"

```

/db_xref="taxon:10090"
/clone="IMAGE:1886966"
/clone_lib="Sugano mouse liver mlia"
/sex="female"

```

```

/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: liver; Vector: pME18S-Fl3; Site_1: DraIII
(CACTCTGCG): site 2: DraIII (CACTATCTG): 1st strand cDNA

```

was primed with an oligo(dT) primer [ATGTGGCCCTTTTTTTTTTTTTT]; double-stranded cDNA was ligated to a DrallI adaptor (TGTGGCCCTACTGc), digested and cloned into distinct DrallI sites of the pMT18S-FR3

vector (5' site CACGTCG, 3' site CACATGTC). *Amo1* should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library

BASE COUNT
ORIGIN

108 a 125 c 111 g 87 t

constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTGTGCTCTAAAGCTGGC
and 3' end primer CGACCTGCACCTGAGACAA.¹⁶

Query Match	3.1%;	Score 152.6;	DB 9;	Length 431;
Best Local Similarity	72.9%;	Pred. No. 6.5e-36;		
Matches 196;	Conservative 0;	Mismatches 73;	Indels 0;	Gaps 0;

95 CAGCACTTGAGCAAAAGCCATGTGGTCCAGGATTGCTACCATGTTGATGACACAGAGTATC 154
 ||| ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 60 CACCAGGAGCAAAACACTGTGGTCCAGGAATGCTACCGAGCGGAGTAGGGGCAGA GTATC 119

9Y 155 GAGGACGTACATCCACCACATGTCACAGGAAGGACCTGCCAAGCTTGGTCATCATATGACAC 214
 | | | | | | | | | | | | | | | | | | | | | |
 Db 120 GGGGTACATCGTCCACATACCATCACAGGAGAAGTGGCAGTCTGGCAGCATATGTTTC 179

213 CAGATCAGCAATATAGGACACAGAGGAGGATACCCAGAAATGCGGCTGGATCATGAGATACCT 274
 180 CACACAGGCATTCGAGAGACCCGAGAGACTTCCAGATGCTGGCTTGGAGATGACACTCT 239

Df

273 GAGGAAATCCAGAAGCAGTGGACGCCTCATTAAGTTAACAAGGAATCCCAGTCACAGGT
||||| | |||| | |||| | |||| | |||||
Db 240 GCAGGAACC CGATTGGTGACCAAGGCCCTTGGTCTCACACTGACC GGCGGTCAAGGT 299

322	GGGAGTATGACCTGGGCGGAAAGCTCA	300
300	GGGATACTGCAACCTGAAGCGGTGCTCA	328

Search completed: March 6, 2003, 15:19:13
Job time : 6007.5 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 01:17:11 ; Search time 12208 seconds

(without alignments)
11919.563 Million cell updates/sec

Title: US-09-923-515-3_COPY_8939_13938

Perfect score: 5000
Sequence: 1 ctcgaccttaccgccggtt.....ctgcactatttgatttga 5000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

GenEmbl: *
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_cm:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pla:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5000	100.0	13938	9 HSAALIPOA	X06290 Human mRNA
2	3747.2	74.9	4510	9 MACCAPOA	J04635 Rhesus monk
3	1237	24.7	2753	6 I45623	I45623 Sequence 1
4	1235.4	24.7	2732	9 HSEPMGR	X05199 Human mRNA
5	1235.4	24.7	2753	6 A22096	A22096 Plasmidogen
6	1232.8	24.7	1784	9 HUMPMGM	K02922 Human Plasm
7	1228.8	24.6	2679	6 I06212	I06212 Sequence 14
8	1178.8	23.6	2711	9 MACPEPSC	J04697 Rhesus monk
9	1110.6	22.2	2296	6 AR105749	AR105749 Sequence
10	1110.6	22.2	2296	6 I14104	I14104 Sequence 18
11	1106.4	22.1	2497	6 AX448883	AX448883 Sequence
12	1104.8	22.1	2497	6 AR037325	AR037325 Sequence
13	1104.8	22.1	2497	6 AR082437	AR082437 Sequence
14	1104.8	22.1	2497	6 AR085163	AR085163 Sequence
15	1104.8	22.1	2497	9 HUMPLASM	M74220 Human Plasm
16	1089.8	21.8	2433	6 AX463622	AX463622 Sequence
17	989.4	19.8	1649	9 AF029691	AF029691 Papio ham
18	880.2	17.6	1272	9 AF029692	AF029692 Papio ham
19	868.6	17.4	2770	4 BTPPLASMIN	X79402 B. taurus mk
20	858.6	17.2	2041	9 HSU19517	HSU19517 Human (apoa
21	857.6	17.2	2771	10 BC014773	BC014773 Mus muscu
22	851.8	17.0	2720	10 MUSEPLGN	J04766 Mouse plasm
23	845.4	16.9	1047	6 AX463626	AX463626 Sequence
24	830.8	16.6	2737	10 RNO242649	EU33171 E. coli
25	775.4	15.5	2745	4 EU33171	EU33171 E. coli
26	740.8	14.8	1712	9 HSU19518	HSU19518 Human (apoa
27	693	13.9	2832	4 AF012297	AF012297 Macropus
28	613	12.3	750	6 AX463624	AX463624 Sequence
29	557.8	11.2	690	6 A37078	A37078 Sequence 1
30	557.8	11.2	690	6 I51668	I51668 Sequence 1
31	410	8.2	135751	6 AL596089	AL596089 Human DNA
32	408.4	8.2	271762	2 AL645523	AL645523 Homo sapi
33	325.4	6.5	171363	2 AC084863	AC084863 Papio cyn
34	311.6	6.2	1353	6 AX117488	AX117488 Sequence
35	306.4	6.1	1236	6 AX329649	AX329649 Sequence
36	306.4	6.1	1236	6 AX336098	AX336098 Sequence
37	306.4	6.1	1236	6 AX336421	AX336421 Sequence
38	306.4	6.1	1236	6 HUMPLG24	M34276 Human Plasm
39	306.4	6.1	2589	9 AK096980	AK096980 Homo sapi
40	306.4	6.1	162083	9 AL591069	AL591069 Human DNA
41	306.4	6.1	173758	2 HSDJ81D8	AC010893 Homo sapi
42	306.4	6.1	173758	2 AC040893	AC040893 Homo sapi
43	306.4	6.1	200402	2 AL645523	AL645523 Homo sapi
44	306.4	6.1	271762	2 AL645523	AL645523 Homo sapi
45	283	5.7	435	9 S79621	S79621 apolipoprot

ALIGNMENTS

RESULT 1
LOCUS HSAALIPOA 13938 bp mRNA linear PRI 30-MAR-1995
DEFINITION Human mRNA for apolipoprotein(a).
ACCESSION X06290.1 M17399 X06966
VERSION X06290.1 GI:28619
KEYWORDS apolipoprotein; apolipoprotein A; glycoprotein; lipoprotein;
plasma; serine protease.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 13938)
AUTHORS McLean, J.W.
TITLE Direct Submission

QY 361 TTCCAGGCTTAGAGGCTCTTCCGAACAGACACGACTGAGGAGGCTTG6GGTCCAG 420
|||||
Db 9299 TTCCAGGCTTAGAGGCTCTTCCGAACAGACACGACTGAGGAGGCTTG6GGTCCAG 9358
QY 421 AGTGTACACGAGTATGACAGAGTATGAGGACACATACCTACCACTGTCACTGGAA 480
|||||
Db 9359 AGTGTACACGAGTATGACAGAGTATGAGGACACATACCTACCTGTCACTGGAA 9418
QY 481 GAACCTGCAAGCTTGGTCAATGACACACACCTGATAGTGGAGCCCAAGATCT 540
|||||
Db 9419 GAACCTGCAAGCTTGGTCAATGACACACACCTGATAGTGGAGCCCAAGATCT 9478
QY 541 ACCCAATGCTGGCTTGATCATGAACATCTGACGAAATCCAGATGCTGTGGAGCTCT 600
|||||
Db 9479 ACCCAATGCTGGCTTGATCATGAACATCTGACGAAATCCAGATGCTGTGGAGCTCT 9538
QY 601 ATTGTATACAGAGGATCCGGGTGTACGTGGGAGTACAGCAACCTGACGCAATGCTAG 660
|||||
Db 9539 ATTGTATACAGAGGATCCGGGTGTACGTGGGAGTACAGCAACCTGACGCAATGCTAG 9598
QY 661 ACCGAGAGGAGCTGCGGTGCGGCTCCGACTGTACCCCGGTTCCAAAGCTTAGAGCTC 720
|||||
Db 9599 ACCGAGAGGAGCTGCGGTGCGGCTCCGACTGTACCCCGGTTCCAAAGCTTAGAGCTC 9658
QY 721 CTTCCGAACAGCAACCGACAGAGAGGCGCTGGGGGAGGAGTCTACCAAGGTAATG 780
|||||
Db 9659 CTTCCGAACAGCAACCGACAGAGAGGCGCTGGGGGAGGAGTCTACCAAGGTAATG 9718
QY 781 GACAGAGTTATCGAGGACATACCTCCACACTGTCACTGGAAAGAACCTCCAAAGCTTGT 840
|||||
Db 9719 GACAGAGTTATCGAGGACATACCTCCACACTGTCACTGGAAAGAACCTCCAAAGCTTGT 9778
QY 841 CATCTATACACCCACACTCGCATAGTGGAGCCCAAGAACTCTCCAAATGCTGGTTGA 900
|||||
Db 9779 CATCTATACACCCACACTCGCATAGTGGAGCCCAAGAACTCTCCAAATGCTGGTTGA 9838
QY 901 TCATGAACTACTGAGGAATCCAGATGCTGTGGCAAGCTCTCTTATTTATACGAGGATC 960
|||||
Db 9839 TCATGAACTACTGAGGAATCCAGATGCTGTGGCAAGCTCTCTTATTTATACGAGGATC 9898
QY 961 CCGGTGAGGAGGAGAGTACGCAACTGACGCAATGCTCAGACGAGAGAGAGCTGCCG 1020
|||||
Db 9899 CCGGTGAGGAGGAGAGTACGCAACTGACGCAATGCTCAGACGAGAGAGAGCTGCCG 9958
QY 1021 TCGCGCTCCGACTGTACCCCGGTTCCAAAGCTAGAGGCTCTCTCCGAACAAGCACGA 1080
|||||
Db 9959 TCGCGCTCCGACTGTACCCCGGTTCCAAAGCTAGAGGCTCTCTCCGAACAAGCACGA 10018
QY 1081 CTGAGCAGAGGCTGGGGTGCAGAGAGTACCAAGGCTATGAGACAGATTCAGAGGCA 1140
|||||
Db 10019 CTGAGCAGAGGCTGGGGTGCAGAGAGTACCAAGGCTATGAGACAGATTCAGAGGCA 10078
QY 1141 CATACTCCACCACTGCTACTGGAAGAACCTGCCAAGCTTGTCATCTATGACACCACT 1200
|||||
Db 10079 CATACTCCACCACTGCTACTGGAAGAACCTGCCAAGCTTGTCATCTATGACACCACT 10138
QY 1201 GCGATAGTCCGACCCCAATATCAACCAATGCTGCTGATCAATGAGTACTGAGAGA 1260
|||||
Db 10139 GCGATAGTCCGACCCCAATATCAACCAATGCTGCTGATCAATGAGTACTGAGAGA 10198
QY 1261 ATCCAGATCTGTGGCAGCCCTTATTTATATACGAGGATCCAGTGTCCAGGTGGAGT 1320
|||||
Db 10199 ATCCAGATCTGTGGCAGCCCTTATTTATATACGAGGATCCAGTGTCCAGGTGGAGT 10258
QY 1321 ACTGGAACCTGACAAAGTCTCAGACGAGAGGAGTGGCGGTCCCAACTATTA 1380
|||||
Db 10259 ACTGGAACCTGACAAAGTCTCAGACGAGAGGAGTGGCGGTCCCAACTATTA 10318
QY 1381 CCCCAGTTTCCAAAGCTTAGAGGCTCTTCTGAAACAGCACCACTGAGCAAAAGGCTGGGG 1440
|||||
Db 10319 CCCCAGTTTCCAAAGCTTAGAGGCTCTTCTGAAACAGCACCACTGAGCAAAAGGCTGGGG 10378

QY 1441 TCGAGAGTGTACACCGGAAATGACAGAGTATCAAGGCACATCTTCACTACTGTCA 1500
|||||
Db 10379 TCGAGAGTGTACACCGGAAATGACAGAGTATCAAGGCACATCTTCACTACTGTCA 10438
QY 1501 CAGGAAGAACTGCGACAGCTTGCTATCTATGACACCACTGCGATATGCGAGCCCGAG 1560
|||||
Db 10439 CAGGAAGAACTGCGACAGCTTGCTATCTATGACACCACTGCGATATGCGAGCCCGAG 10498
QY 1561 CATACTACCCAAATGCTGGCTTGATCAAGAACTACTGCGAAATCCAGATCTGTGGCAG 1620
|||||
Db 10499 CATACTACCCAAATGCTGGCTTGATCAAGAACTACTGCGAAATCCAGATCTGTGGCAG 10558
QY 1621 CCCCTTGCTGTATACAAAGATCCCAAGTCCAGTGTGGAGGAGTCTGACACTGACAGT 1680
|||||
Db 10559 CCCCTTGCTGTATACAAAGATCCCAAGTCCAGTGTGGAGGAGTCTGACACTGACAGT 10618
QY 1681 GCTCAGATCAGAAATGAGTGCCTCGCTCCGTAATGTTATCTGGCTCCAAAGCTAG 1740
|||||
Db 10619 GCTCAGATCAGAAATGAGTGCCTCGCTCCGTAATGTTATCTGGCTCCAAAGCTAG 10678
QY 1741 AGGCTTTTATTAACAACAAGCACTGACTAGAGAAACCCCGGGGTACAGACTGCTACTAC 1800
|||||
Db 10679 AGGCTTTTATTAACAACAAGCACTGACTAGAGAAACCCCGGGGTACAGACTGCTACTAC 10738
QY 1801 ATATGACAGAGTTAACGAGGCACTATCCACCACTGTCACAGAGAACTTGGCAAG 1860
|||||
Db 10739 ATATGACAGAGTTAACGAGGCACTATCCACCACTGTCACAGAGAACTTGGCAAG 10798
QY 1861 CTTGTGATCTATGACACCACACAGCATAGTCCGAGCCCGAAACTACCAATGCTG 1920
|||||
Db 10799 CTTGTGATCTATGACACCACACAGCATAGTCCGAGCCCGAAACTACCAATGCTG 10858
QY 1921 GCCTGACAGAACTACTGACAGAAATCCAGATGCTGGAATGCGCTTGCTGTACACCA 1980
|||||
Db 10859 GCCTGACAGAACTACTGACAGAAATCCAGATGCTGGAATGCGCTTGCTGTACACCA 10918
QY 1981 TGGATCCCAAGTGCAGGTGGAGTACTCAACTGACCAATGCTGTGGTACAGAAATCAA 2040
|||||
Db 10919 TGGATCCCAAGTGCAGGTGGAGTACTCAACTGACCAATGCTGTGGTACAGAAATCAA 10978
QY 2041 GTGTCTTGCAACTCTCACGGTGTGCCAGATCCAAAGCACAGAGGCTCTTCTGAAAGAG 2100
|||||
Db 10979 GTGTCTTGCAACTCTCACGGTGTGCCAGATCCAAAGCACAGAGGCTCTTCTGAAAGAG 11038
QY 2101 CACCAACGAGCAAGCCCGGGTCCAGAGATTGCTACATGGTGTATGACAAAGTTATC 2160
|||||
Db 11039 CACCAACGAGCAAGCCCGGGTCCAGAGATTGCTACATGGTGTATGACAAAGTTATC 11098
QY 2161 GAGGCTCATCTCTACCACTGTCACAGAGAGACATGTCAGTCTTGGCTCTATGACAC 2220
|||||
Db 11099 GAGGCTCATCTCTACCACTGTCACAGAGAGACATGTCAGTCTTGGCTCTATGACAC 11158
QY 2221 CACACTGCACTCAGAGAGACACAGAAATTTATCCAAATGCTGGCTTACACAGAACTACT 2280
|||||
Db 11159 CACACTGCACTCAGAGAGACACAGAAATTTATCCAAATGCTGGCTTACACAGAACTACT 11218
QY 2281 GCAAGAAATCCAGATGCTGATTAAGTCTGGTGTATATACCATGATCCCAATGTCAGAT 2340
|||||
Db 11219 GCAAGAAATCCAGATGCTGATTAAGTCTGGTGTATATACCATGATCCCAATGTCAGAT 11278
QY 2341 GGGAGTACTGCAACTGACACAATGTCAGTGTACAGAAATCAAGTGTCTTGGCAGCTCCA 2400
|||||
Db 11279 GGGAGTACTGCAACTGACACAATGTCAGTGTACAGAAATCAAGTGTCTTGGCAGCTCCA 11338
QY 2401 CCGCTGTTTCTGAAACAAGCAACCAAGGAGCAAAAGCCCAAGTCCAGAGCTGTACACTG 2460
|||||
Db 11339 CCGCTGTTTCTGAAACAAGCAACCAAGGAGCAAAAGCCCAAGTCCAGAGCTGTACACTG 11398
QY 2461 GTGATGACAGAGTTATCGAGGCTATCTCCACCACTGTTACAGAGAAAGATGTGAGT 2520
|||||
Db 11399 GTGATGACAGAGTTATCGAGGCTATCTCCACCACTGTTACAGAGAAAGATGTGAGT 11458
QY 2521 CTTGGTCTCTATGACACCACTGTCAGAGAAACCAAGAAATACTACCAAAATGCTG 2580

Db 11459 CTTCGCTCTCTATGACACCCACATGGCATCAGAGAAACCCAGAAATCTATCCCAATGCTG 11518
Qy 2381 GCTTGACCCAGAACTACTGAGAGATTCAGATGCTGAGATTCGCCCTTGGTGTATATACA 2640
Db 11519 GCCGACCCAGAACTACTGAGAGATTCAGATGCTGAGATTCGCCCTTGGTGTATATACA 11578
Qy 2641 TGGATCCCACTGTCAGATGGAGATCTGCACTGACCGAATGTCACATGATGATGATCAA 2700
Db 11579 TGGATCCCACTGTCAGATGGAGATCTGCACTGACCGAATGTCACATGATGATGATCAA 11638
Qy 2701 CTCTCTCACAACCTCCAGGGGTGCTCCAGTTCCAGACACAGAGCTTCCTTCTGAGAAG 2760
Db 11639 CTCTCTCTCACAACCTCCAGGGGTGCTCCAGTTCCAGACACAGAGCTTCCTTCTGAGAAG 11698
Qy 2761 CACCACTGAAACAGACACTGGGGTCCAGAGCTGCTACCGAGTGTATGAGAGTTATC 2820
Db 11699 CACCACTGAAACAGACACTGGGGTCCAGAGCTGCTACCGAGTGTATGAGAGTTATC 11758
Qy 2821 GAGGCACACTCTCCACACTATTCACAGAAACATGTCAGTCTTGGTCTCTATGACAC 2880
Db 11759 GAGGCACACTCTCCACACTATTCACAGAAACATGTCAGTCTTGGTCTCTATGACAC 11818
Qy 2881 CACATTTGCGATGAGAGATCCCATTTATCTATCCAAATGCTGGCCCTGACCGAATCT 2940
Db 11819 CACATTTGCGATGAGAGATCCCATTTATCTATCCAAATGCTGGCCCTGACCGAATCT 11878
Qy 2941 GCAGGAATCCAGATGCTGAGATTCGCCCTTGGTGTATACCATGATGCCAGTGTACAGT 3000
Db 11879 GCAGGAATCCAGATGCTGAGATTCGCCCTTGGTGTATACCATGATGCCAGTGTACAGT 11938
Qy 3001 GGGAGTCTGCTCACTGACACGATGTCAGTACAGAAATCGAGTGTCTCTCACTCTCCA 3060
Db 11939 GGGAGTCTGCTCACTGACACGATGTCAGTACAGAAATCGAGTGTCTCTCACTCTCCA 11998
Qy 3061 CAGTGGCCCCGGTTCACAGACAGAGGCTCTTCTGAAACAAAGCACACCTGAGAAAAACC 3120
Db 11999 CAGTGGCCCCGGTTCACAGACAGAGGCTCTTCTGAAACAAAGCACACCTGAGAAAAACC 12058
Qy 3121 CTGTGTGTCAGAGATGCTCTACCATGATGATGAGAGGATTTACGAGGCTATCTCCACCA 3180
Db 12059 CTGTGTGTCAGAGATGCTCTACCATGATGATGAGAGGATTTACGAGGCTATCTCCACCA 12118
Qy 3181 CTGTCTACAGAGAGACCTGTCAATCTTGGTCACTCTATATACACACACTGCGCATGAGA 3240
Db 12119 CTGTCTACAGAGAGACCTGTCAATCTTGGTCACTCTATATACACACTGCGCATGAGA 12178
Qy 3241 CCCCAGAAAATCTACCAATGCTGCGCTGACCGAAGTACTGACAGAAATCCAGATTCTG 3300
Db 12179 CCCCAGAAAATCTACCAATGCTGCGCTGACCGAAGTACTGACAGAAATCCAGATTCTG 12238
Qy 3301 GGAACAAACCTTGGTGTTCACACACCATCGTGTGAGTGGAGTACTGCAATCTGA 3360
Db 12239 GGAACAAACCTTGGTGTTCACACACCATCGTGTGAGTGGAGTACTGCAATCTGA 12298
Qy 3361 CACAATGCTCAGAAACAGATAGGTGTCCTAGAGACTCCACATGTTGTTCAATCTCAA 3420
Db 12299 CACAATGCTCAGAAACAGATAGGTGTCCTAGAGACTCCACATGTTGTTCAATCTCAA 12358
Qy 3421 GCATGGAAGCTCATTTCTGAAGACGACCAACTGACAAACCCCTGTGTCGGCAGTGTCT 3480
Db 12359 GCATGGAAGCTCATTTCTGAAGACGACCAACTGACAAACCCCTGTGTCGGCAGTGTCT 12418
Qy 3481 ACCATGGTATGCGCAGAGTTATCGAGGCACATTTCTCCACACTGTGTCACAGAGAGCAT 3540
Db 12419 ACCATGGTATGCGCAGAGTTATCGAGGCACATTTCTCCACACTGTGTCACAGAGAGCAT 12478
Qy 3541 GTCAATCTTGTGTCATGACAGACACACCGGATCAGAGACCCCGAAGAACTATCCCAA 3600
Db 12479 GTCAATCTTGTGTCATGACAGACACACCGGATCAGAGACCCCGAAGAACTATCCCAA 12538
Qy 3601 ATGATGGCTGACATGAACTACTGACAGAAATCCAGATGCGGATACAGGCCCTTGGTGT 3660

Db 12539 ATGATGGCTGACATGAACTACTGACAGAAATCCAGATGCCATACAGGCCCTTGGTGT 12598
Qy 3661 TTACCATGGACCCAGACATCAGTGGGATGTTCTGCAACTGACCGGATGCTCAGACACAG 3720
Db 12599 TTACCATGGACCCAGACATCAGTGGGATGTTCTGCAACTGACCGGATGCTCAGACACAG 12658
Qy 3721 AAGGACTGTGTGCTGCTCCCGACTGCTCATCCAGTTCGAAGCCGAGAGGCCCTCTTGTG 3780
Db 12659 AAGGACTGTGTGCTGCTCCCGACTGCTCATCCAGTTCGAAGCCGAGAGGCCCTCTTGTG 12718
Qy 3781 AACAGACTGTATGTTTGGGATGGAAGATACCGGGGCAAGAAAGCAACCACTGTTA 3840
Db 12719 AACAGACTGTATGTTTGGGATGGAAGATACCGGGGCAAGAAAGCAACCACTGTTA 12778
Qy 3841 CTGGAGCGCATGCGCAGGAATGGGCTCCGAGAGGCCCATGACACAGCAGCTTCATTC 3900
Db 12779 CTGGAGCGCATGCGCAGGAATGGGCTCCGAGAGGCCCATGACACAGCAGCTTCATTC 12838
Qy 3901 CAGGACAAATTAATGAGGCTCTGGAATAAAATTAATCTGCCGTAAACCTGATGGTGACA 3960
Db 12839 CAGGACAAATTAATGAGGCTCTGGAATAAAATTAATCTGCCGTAAACCTGATGGTGACA 12898
Qy 3961 TCAATGTCCTGCTGCTACACAAATGAATCCAGAAACTTTTGTGACTGCTGATATCC 4020
Db 12899 TCAATGTCCTGCTGCTACACAAATGAATCCAGAAACTTTTGTGACTGCTGATATCC 12958
Qy 4021 CTCTCTGTGATCTCTTCAATTTGATTTGTGGGAAGCCTCAAGTGAAGCCGAGAAATGTC 4080
Db 12959 CTCTCTGTGATCTCTTCAATTTGATTTGTGGGAAGCCTCAAGTGAAGCCTGATATCC 13018
Qy 4081 CTGGAGCATTTGTGGGGGGGTGTGGGCCACCAATTCCTGGCCCTGGCAAGTCAATC 4140
Db 13019 CTGGAGCATTTGTGGGGGGGTGTGGGCCACCAATTCCTGGCCCTGGCAAGTCAATC 13078
Qy 4141 TCAGAACAGGTTTGGAAACACCTCTGTGAGGACCTTAATATTCACAGATGGGTGC 4200
Db 13079 TCAGAACAGGTTTGGAAACACCTCTGTGAGGACCTTAATATTCACAGATGGGTGC 13138
Qy 4201 TGACGTGCTGCTCATCTGCTTGAAGAGCTCAAGGCCCTTCATCTACAAAGGCTATCCGG 4260
Db 13139 TGACGTGCTGCTCATCTGCTTGAAGAGCTCAAGGCCCTTCATCTACAAAGGCTATCCGG 13198
Qy 4261 GTGCACACCAAGAGTGAACCTCGAATCTCATGTTAGGAAATTAAGTGTCTAGGCTGT 4320
Db 13199 GTGCACACCAAGAGTGAACCTCGAATCTCATGTTAGGAAATTAAGTGTCTAGGCTGT 13258
Qy 4321 TCTTGGAGCCACACAGCAGATATTGCTGCTTAAGGTTAAGCAGGCCCTGCCATCA 4380
Db 13259 TCTTGGAGCCACACAGCAGATATTGCTGCTGCTTAAGGTTAAGCAGGCCCTGCCATCA 13318
Qy 4381 CTGACAAAGTAAATCCAGCTTGTCTGCCATCCCCAGACTCATGCTACCCGACAGACTG 4440
Db 13319 CTGACAAAGTAAATCCAGCTTGTCTGCCATCCCCAGACTCATGCTACCCGACAGACTG 13378
Qy 4441 AATGTTACATCACTGCTGGGGAAGAAACCAAGGTCCTTTGGACTGCGCTTCTCAAG 4500
Db 13379 AATGTTACATCACTGCTGGGGAAGAAACCAAGGTCCTTTGGACTGCGCTTCTCAAG 13438
Qy 4501 AAGCCACGCTCTTGTATTGAGAAATGAAGTGTGCAATCACTATTAAGTATTTGTGCTG 4560
Db 13439 AAGCCACGCTCTTGTATTGAGAAATGAAGTGTGCAATCACTATTAAGTATTTGTGCTG 13498
Qy 4561 AGCAATTTGGCAGAGCAGTGAAGTGTGCAAGGTTGACAGTGAAGGCTCTGTTGTTGCT 4620
Db 13499 AGCAATTTGGCAGAGCAGTGAAGTGTGCAAGGTTGACAGTGAAGGCTCTGTTGTTGCT 13558
Qy 4621 TCGAAGAGACAAATACATTTTCAAGAGTCACTTCTTGGGCTTGTGGCTGTGACGCC 4680
Db 13559 TCGAAGAGACAAATACATTTTCAAGAGTCACTTCTTGGGCTTGTGGCTGTGACGCC 13618
Qy 4681 CCAATTAAGCCTGTGCTATGCTGCTGTTCAAGGTTTGTACTTGAATGAGGAATGA 4740
Db 13619 CCAATTAAGCCTGTGCTATGCTGCTGTTCAAGGTTTGTACTTGAATGAGGAATGA 13678

QY 1176 GCTTGCTATCTATGACACACACTCGCATAGTGGACCCGAGAAATCTACCCAAATGCT 1235
|||||
Db 901 GCTTGGTCTATGTAAGACACACACTCTCATAGTGGACCCGAGAAATCTACCCAAATGCT 960
QY 1236 GCGTGTATGATTAAGTACTGAGAGAAATCCAGATCTGTGGACGCCCTTATGTAATGAG 1295
|||||
Db 961 GCGTGTATGATTAAGTACTGAGAGAAATCCAGATCTGTGGACGCCCTTATGTAATGAG 1020
QY 1296 AGGAGATCCAGTGTAGAGTGGAGTACTGACACTGACAAATGCTCAGACAGAGAAAGG 1355
|||||
Db 1021 ATGGATCCCAATGATGAGGAGAGTACTGACAACTGACAAATGCTCAGACAGAGAAAGG 1080
QY 1356 ACTGCGGTGCGCTCCAACTATTACCCGATTCGAAGCCCTAGAGGCTCTTCTGAAACAA 1415
|||||
Db 1081 ACTGCGGTGCGCTCCCAATGATGACCCGGTTCGAAGCCCTAGAGGCTCTTCTGAGCAA 1140
QY 1416 GGCACAACTGACAAAGGCTGGGGTGCAGAGTGCATCCACAGAAATGACAGACTTAT 1475
|||||
Db 1141 GGCACAACTGACAAAGGCTGGGGTGCAGAGTGCATCCACAGAAATGACAGACTTAT 1200
QY 1476 CAAGGACAACTCTATTAAGTGCACAGAGAAACCTGCGAAGCTTGGTCAATCTATGACA 1535
|||||
Db 1201 CGAGGACAACTACTTACACACTGTGACAGAGAAACCTGCGAAGCTTGGTCAATCTATGACA 1260
QY 1536 CCACACTGCAATAGTGCAGCCGACAGCACTACTACCAATGCTGGCTTATCAAGACTAC 1595
|||||
Db 1261 CCACACTGCAATAGTGCAGCCGACAGCAACCTACTACCAATGCTGGCTTATCAAGACTAC 1320
QY 1596 TGGCGAAATCCAGATCTGTGGACGCCCTTGGTGTATACAAAGATCCCACTGTCAAG 1655
|||||
Db 1321 TGGCGAAATCCAGATCTGTGGACGCCCTTGGTGTATACAAAGATCCCACTGTCAAG 1380
QY 1656 TGGGAGTACTGCAACCTGACAGAGATGCTCAGATCAGAAATGAGTGCCTTCTGTCCTCG 1715
|||||
Db 1381 TGGGAGTACTGCAACCTGACAGAGATGCTCAGATCAGAAAGGACTGCTCATGCTCTCA 1440
QY 1716 AATGATATCTGCTGCAAGGCTTAGAGGCTTTTGAACAACAGCACTGAGAGAAAC 1775
|||||
Db 1441 AATGATATCTGCTGCAAGGCTTAGAGGCTTTTGAACAACAGCACTGAGAGAAAC 1500
QY 1776 CCGGGGTACAGAGACTGCTACTACCAATTATGACAGAGTTACCGAGGACATCTCACCC 1835
|||||
Db 1501 CCGGGGTACAGAGAGTGTCTACTACCAATTATGACAGAGTTATGAGGCACTACTCTCACCC 1560
QY 1836 ACTGTCACAGAGAACTTGCAGGCTGTGATCTATGACACACACAGCACTAGTACG 1895
|||||
Db 1561 ACTGTCACAGAGAACTTGCAGGCTGTGATCTATGACACACACAGCACTAGTACG 1620
QY 1896 ACCCCAGAAATCTACCCAAATGCTGGCTGACACGAACTACTGACAGAAATCCAGATGCT 1955
|||||
Db 1621 ACCCCAGAAATCTATCCAAATGCTGGCTGACACGAACTACTGACAGAAATCCAGATGCT 1680
QY 1956 GAGATTCGCCCTTGGTGTAGACAGATGAGTCCAGTGTCAAGTGGAGTACTGCAACTG 2015
|||||
Db 1681 GAGATTCGCCCTTGGTGTATACCAATGATCCAGTGTCAAGTGGAGTACTGCAACTG 1740
QY 2016 ACCCAATGCTGCTGACAGAAATCAAGTGTCTTGAACAAGTGCAGGAGTCCCAAGTCCA 2075
|||||
Db 1741 ACCCAATGCTGCTGACAGAAATCAAGTGTCTTGAACAAGTGTGCAGGAGTCCCAAGTCCA 1800
QY 2076 AGCAGAGAGCTTCTTCTGAAGAAGCAACAGAGCAAGGCCGCGGGTCCAGAAATTCG 2135
|||||
Db 1801 AGCAGACAGAGCTTCTTCTGAAGAAGCAACAGAGCAAGGCCGCGGGTCCAGAGCTTC 1860
QY 2136 TACCAATGATGATGACAGAAATATGAGGCTCATCTTCCACTGTCACAGAAAGGACA 2195
|||||
Db 1861 TACCAATGATGATGACAGAAATATGAGGCTCATCTTCCACTGTCACAGAAAGGACA 1920
QY 2196 TGTCACTCTTGGTCTCTATGACACACAACTGCAATGAGAGCAACAGAAATATTATCCA 2255
|||||
Db 1921 TGTCACTCTTGGTCTCTATGACACAACTGCAATGAGAGCAACAGAAATATTATCCA 1980
QY 2256 AATGATGCGCTGACAGAGAACTACTGACAGAAATCCAAATGCTGAGATTAAGTCTTGGTGT 2315

Db 1981 GATGATGCGCTGACAGAGAACTACTGACAGAAATCCAAATGCTAGATTTGCCCTTGGTGT 2040
|||||
QY 2316 TATACCATGATCCCAATGATGAGATGAGTACTGCAACTGACACAAATGCTCAAGTACA 2375
|||||
Db 2041 TATACCATGATCCCAATGATGAGTGGAGTACTGCAACTGACACAAATGCTCAAGTACA 2100
QY 2376 GATCAAGTGTCTGCAAGCTCCAGGCTTTTCTTAACAACACACAGAGCAAGCAAGC 2435
|||||
Db 2101 GAATCAAGTGTCTGCAAGCTCCAGGCTTTTCTTAACAACACACAGAGCAAGCAAGC 2160
QY 2436 CCCACAGTCCAGAGTGTACTACATGATGATGACAGAGTATGAGGCTCATATTCTCACCC 2495
|||||
Db 2161 CCCGGGGTCCAGAGTGTACTACATGATGATGACAGAGTATGAGGCTCATATTCTCCACC 2220
QY 2496 ACTGTTACAGAGAGACATGTCAGTCTTGGTCTCTATGACACACAGAGGATCAGAGA 2555
|||||
Db 2221 ACTGTCACAGAGAGATGTCAGTCTTGGTCTCTATGACACACAGAGGATCAGAGG 2280
QY 2556 ACCACAGAAATCTACCCAAATGCTGGCTGACAGAGAACTACTGACAGAAATCCAGATGCT 2615
|||||
Db 2281 ACCATGAAATCTACCCAAATGCTGGCTGACAGAGAACTACTGACAGAAATCCAGATGCT 2340
QY 2616 GAGATTCGCCCTTGGTGTATACCATGATCCAGTGTCAAGTGGAGTACTGCAACTG 2675
|||||
Db 2341 GAGATTCGCCCTTGGTGTATACCATGATCCAGTGTCAAGTGGAGTACTGCAACTG 2400
QY 2676 AGCAATGTCAGTATGAGATCAACTCTCTCAAACTCCCAAGGCTGCCCATGCTCA 2735
|||||
Db 2401 ACCAATGTCAGTATGAGATCAAGTGTCTTCAAACTCCCAAGGCTGCCCATGCTCA 2460
QY 2736 AGCAGAGACTCTCTCTGAAGAAGCAACCACTGAAACAGCACTGGGTCCAGAGACTGC 2795
|||||
Db 2461 AGCAGAGAGTCTCTCTGAAGAAGCAACCACTGAAACAGCACTGGGTCCAGAGACTGC 2520
QY 2796 TACCGAGTATGAGAGATTAACAGAGGCACTCTCCAGCACTATCAAGAGAAACA 2855
|||||
Db 2521 TACCAAGTATGAGAGATTAACAGAGGCACTCTCCAGCACTATCAAGAGAAACA 2580
QY 2856 TGTCACTTGGTGTCTATGACACACAAATGCTGGAGATGCCATTAATCTATCCA 2915
|||||
Db 2581 TGTCACTTGGTGTCTATGACACACAAATGCTGGAGATGCCATTAATCTATCCA 2640
QY 2916 AATGCTGSCCTGACAGAGAACTACTGACAGAAATCAGATCTGAGATTTGCCCTTGGTGT 2975
|||||
Db 2641 AATGCTGSCCTGACAGAGAACTACTGACAGAAATCAGATCTGAGATTTGCCCTTGGTGT 2700
QY 2976 TACACCATGATCCCAAGTGTAGAGTGGAGTACTGCACTGACACAGATGCTCAATGACA 3035
|||||
Db 2701 TACACCATGATCCCAAGTGTAGAGTGGAGTACTGCACTGACACAGATGCTCAATGACA 2760
QY 3036 GAATCAGAGTGTCTCTCACAACTCCCAAGTGGCCCGGTTCCAAAGCAGAGAGCTCTCTCT 3095
|||||
Db 2761 GAATCAGAGTGTCTCTCACAACTCCCAAGTGGTCCCGTTCCAAAGCAGAGAGCTCTCTCT 2820
QY 3096 GAACAAAGCAACCTGAGAGAAAGGCCCTGTGGTCCAGGATTTGCTACAGATGATGAGG 3155
|||||
Db 2821 GAACAAAGCAACCTGAGAGAAAGGCCCTGTGGTCCAGGATTTGCTACAGATGATGAGG 2880
QY 3156 AGTTATCGAGGATTCCTCCACAGTGTCAAGAGAAAGCTGTCAATCTTGTCTATCT 3215
|||||
Db 2881 AGTTATCGAGGATTCCTCCACAGTGTCAAGAGAAAGCTGTCAAGTCTTGTCTATCT 2940
QY 3216 ATGATTCACAGACTGCAATCAGAGAGCCGAGAAACCTACCAATGCTGGCTGACGAG 3275
|||||
Db 2941 ATGATTCACAGACTGCAATCAGAGAGCCGAGAAACCTACCAATGCTGGCTGACGAG 3000
QY 3276 AACTACTGAGAGAAATCGAATTCGGAACAAACCTGTGTATCAACAACGATCCGTGT 3335
|||||
Db 3001 AACTACTGAGAGAAATCGAATTCGGAACAAACCTGTGTATCAACAACGATCCAGT 3060
QY 3336 GTGAGTGGAGTACTGAAATCTGACCAATGCTGAGAAACAAATCAGATGCTTCTAGG 3395
|||||

3279 TACTGCAGAAATCCAGATTCTGGGAAACAACCCCTGGTGTATACACAACCGATCCGTGTG 3338
1031 TACTGCCGATCTCTGACGAAAAAGGCCCAATGGTGTGCATACAAACCAAGCCAAAGTG 1090
3339 AGGTGGAGTATCTGACATCTGACACAATGTCTGAAAACGAATCAAGTGTCTAGAGACT 3398
1091 CGGTGGAGTATCTGTAAAGATACCGTCTG----- 1119
3399 CCCATGTGTTCAGATTCACAGATGAGAGCTCATTTGAAAGCAGCAACCAATGAGCAA 3458
1120 -----TGACTCTCCAGATATCCAGCAATTTGGCTCCAGCAACCACTGAGCTA 1174
3459 ACCCTGTGTGCGGAGTGTCTACCATGTGTAATGGCAGAGTATGAGAGCAATTTCTCC 3518
1175 ACCCTGTGTGCGGAGTGTCTACCATGTGTAATGGCAGAGTATGAGAGCAATTTCTCC 1234
3519 ACCATGTGCAGAGAAAGATATGTAATCTTGTGTATCCATGACACCAACCGGATCAG 3578
1235 ACCACACACCCAGAAAGAGTGTGAGTGTGCTATCTATACACCAACCGGATCAG 1294
3579 AGGACCCAGAAATCTACCCAAATGATGGCTGACAAATGAACTACTGAGAAATCCAGAT 3638
1295 AAGACCCAGAAATCTACCCAAATGATGGCTGACAAATGAACTACTGAGAAATCCAGAT 1354
3639 GCCGATACAGGCTTGTGTGTATACATGACCCAGCATGAGTGGAGATGAGTGCAGAC 3698
1355 GCCGATACAGGCTTGTGTGTATACATGACCCAGCATGAGTGGAGATGAGTGCAGAC 1414
3699 CTGACGAGATCTGACACAGAAAGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3758
1415 CTGAAAAATGCTGACGAAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1474
3759 CCAAGCCTAGGAGCTTGT 3818
1475 CCAAGT 1534
3819 GCGAAGAAAGCAACCATGT 3878
1535 GCGAAGAAAGCAACCATGT 1594
3879 CATAGACACAGCTTGT 3938
1595 CATAGACACAGCTTGT 1654
3939 TGGCGTAACTGT 3998
1655 TGGCGTAACTGT 1714
3999 CTCTTGT 4058
1715 CTCTTGT 1774
4059 CAAAGTGGAGCGGAAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4118
1775 CAAAGTGGAGCGGAAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1834
4119 TCTGTGGCTTGT 4178
1835 TCTGTGGCTTGT 1894
4179 TTAATATCCAGAGT 4238
1895 TTAATATCCAGAGT 1954
4239 TCAATCTCAAGATGT 4298
1955 TCAATCTCAAGATGT 2014
4299 GAAATAGAGT 4358
2015 GAAATAGAGT 2074

4359 CTAAAGAGGCTGCGCTGATCTGACAAAGTAATGCCAGCTTGTGTGTGTGTGTGTGTGT 4418
2075 CTAAAGAGGCTGCGCTGATCTGACAAAGTAATGCCAGCTTGTGTGTGTGTGTGTGTGT 2134
4419 TACATGT 4478
2135 TATGT 2194
4479 TTTGT 4538
2195 TTTGT 2254
4539 CACTATAGT-----ATATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4571
2255 CGCTATAGT 2314
4572 AGAGCACTGACAGT 4631
2315 GAGGCACTGACAGT 2374
4632 AAATATATTTTACAAAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4691
2375 AAATATATTTTACAAAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2434
4692 GGT 4751
2435 GGT 2494
4752 TAAATGT 4811
2495 TAAATGT 2554
4812 ATTTAGCATGT 4871
2555 ATTTAGCATGT 2614
4872 CTATGCGCAACCTGT 4931
2615 CTATGCGCAACCTGT 2657
4932 AAATCTGTATTAAGT 4991
2658 GATTTCTGTATTAAGT 2717
4992 TTTGATTT 4998
2718 TTTGATTT 2724

RESULT 4
HSPMGR
LOCUS 2732 bp mRNA linear PRI 21-MAR-1995
DEFINITION Human mRNA for plasminogen.
ACCESSION X05199
VERSION X05199.1 GI:35530
KEYWORDS plasminogen; signal peptide.
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2732)
Forsgren,M., Reden,B., Israelsson,M., Larsson,K. and Heden,L.O.
TITLE Molecular cloning and characterization of a full-length cDNA clone
for human plasminogen
FEBS Lett. 213 (2), 254-260 (1987)
JOURNAL
MEDLINE 87162490
PUBMED 3030813
COMMENT Data kindly reviewed (21-DEC-1987) by HEDEN L.-O.
FEATURES
location
1..2732
/organism="Homo sapiens"
/db_xref="taxon:9606"
55..2487
CDS

```
/note="precursor polypeptide (AA -19 to 791)"
/codon_start=-1
/protein_id="CAA28831.1"
/db_xref="GI:35531"
/db_xref="Swiss-Prot:P00747"
/translation="MEHKEVILLILLFLKSGGPELDYVNTNGASLSPYVKKQLGAG
SIEEACKEDEEPEFCRAFOYHKKRQDQVIAENRKSIIIRMDVYLPERKYLSE
CTKGNKNTGYSKTRKNGITCOKMSSTPHPRSPATHSEGLEBYENYCNPNRDO
GPMVYTTDPERKYDYLILSEBECMHGSEYDGRISKTSGLSCQAMDSQSHAG
YIPSEKPNKMLKKNYCNRPDLRPMCFETTDNKKWELCDIIPCTTPPSSGPTQCL
KGTENYRGNAVAVTSVGHCOHWSAQTPTNHRTEPNEPCRLDENTCRNDGKRAM
CHTNGSREMEYCKIPSCDSPEYTSQALPAPPLTFYVDDCYGDDQSVRGSTST
TTKKCOSMSMPPHROKTPENYPMAGLIMNYCNRPADGKPMCFETTDPERYENCN
LKKSQTEASVMAPPVYLLPVEYTPSEEDCMGNGKGRKGRATYTGTCODMAAO
EPHRSILFPEPTNPRAKLEKNCRNPDGPKYVNTNPKRKYDQVPOCAAASPD
CGKPOVEPKCPGRVYVGGCYAHHSPWQVSIKRTFGHFECCGTLISEBYLITARCL
EKSPRSPSYVILGAHOEVNLEPHVDELVSLEPTKRDILKLSIPATIDKYL
PACLPSPNYVADRTCEFTTGWGTGTGAGLLEADLPYIENKRNRYEFLNGRVO
STELCAGHLAGGTDSDGSDGSPYVCFEKDYILOGVTSWGLGCAKPNKPGVYVRSR
FVWIEGVMMNN"
55.111
112.2484
/product="put. mature plasminogen (AA 1-791)"
/misc_feature
1792.1797
/note="cleavage site"
2689.2694
/misc_feature
/note="pot. polyA signal"
2732
/polyA_site
/note="polyA site"
2732
BASE COUNT 758 a 669 c 688 g 617 t
ORIGIN
Query Match 24.78; Score 1235.4; DB 9; Length 2732;
Best Local Similarity 80.98; Pred. No. 0;
Matches 1511; Conservative 0; Mismatches 276; Indels 80; Gaps 3;
```

```
1405 CTGAAAAAATGCTCAGAGAACGAGACGGTGTGTACACCTCGCCCTGTGCTCTGCTT 1464
1406 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3759 CCAGAGCTAGGAGCCCTCTTTCAGAACAGACTGTATGTTGGGAATGGGAAGGATACCGG 3818
1465 CCAGATGTAGAGACTCTCCGAGAGAAAGACTGTATGTTGGGAATGGGAAGGATACCGA 1524
3819 GCGAAGAAAGCAACACACTGTACTGTGGACGCCATGCCAGAAAGGGCTGCCAGAGGCC 3878
1525 GCGAAGAGGCGACACACTGTACTGTGGACGCCATGCCAGAAAGGGCTGCCAGAGGCC 1584
3879 CATAGACACAGACGCTGATTCACAGGACCAATTAATGAGCAGAGTGTGAAAAAATATAC 3938
1585 CATAGACACAGACGCTGATTCACAGGACCAATTAATGAGCAGAGTGTGAAAAAATATAC 1644
3939 TGGCGTAAACCTTGATGTGACATCAATGTGTCCCGGTGTATACCAATGATGATCAAGAAA 3998
1645 TGCCGTAAACCTTGATGTGACATCAATGTGTCCCGGTGTATACCAATGATGATCAAGAAA 1704
3999 CTTTGTGACTAGTGTGATATCCCTCTGTGATCATCTTCAATTTGATGGAGGCT 4058
1705 CTTTGTGACTAGTGTGATATCCCTCTGTGATCATCTTCAATTTGATGGAGGCT 1764
4059 CAGGTGAGCCGCAAGAAATGTCTGGAAGCATTTGAGGGGGGTGTGGCCACCCACAT 4118
1765 CAGGTGAGCCGCAAGAAATGTCTGGAAGCATTTGAGGGGGGTGTGGCCACCCACAT 1824
4119 TCCTGGCCCTGGCAAGTCACTGTCAGAACAGATTGGAAGACACTTCTGTGAGGACACC 4178
1825 TCCTGGCCCTGGCAAGTCACTGTCAGAACAGATTGGAAGACACTTCTGTGAGGACACC 1884
4179 TTAATATCCGACAGTGGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4238
1885 TTAATATCCGACAGTGGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1944
4239 TCATCTCAAGAGTCACTCTGAGTGCACACCAAGAGTGAACCTCGAATCTCATGTTTCAG 4298
1945 TCATCTCAAGAGTCACTCTGAGTGCACACCAAGAGTGAACCTCGAATCTCATGTTTCAG 2004
4299 GAAATGAAAGTGTCTGAGTGTCTGAGGACCAACAGACGATTTGCTGTCAAG 4358
2005 GAAATGAAAGTGTCTGAGTGTCTGAGGACCAACAGACGATTTGCTGTCAAG 2064
4359 CTAGACAGGCTGCGCTCATCTGACAGCAAGTAATGCCAGCTGTGCTGCATCCAGAC 4418
2065 CTAGACAGGCTGCGCTCATCTGACAGCAAGTAATGCCAGCTGTGCTGCATCCCAAT 2124
4419 TACATGTCACGCCAGACGATGATTTACATCACTGCTGGGAGAGAACCCAGAGTACC 4478
2125 TATGTGTGCTGCTGACCGAGCCGAAATGTTCACTGGCTGTGGAGAGAACCCAGAGTACC 2184
4479 TTTGGGACTGGCCCTTCTCAAGAGCCCAAGCTCCTGTGATGAGTAAGTGTCAAT 4538
2185 TTTGGGACTGGCCCTTCTCAAGAGCCCAAGCTCCTGTGATGAGTAAGTGTCAAT 2244
4539 CACTATAAGT-----ATAATTTGCTGAGCAATTTGGCC 4571
2245 CGCTATGAGTGTGTAATGAGAGTCCATCCACCCAGACCTGTGTGGGCAATTTGGCC 2304
4572 AGAGGCACTGACAGTTGCCAGGCTGACAGTGAAGGCTCTGTGTTTCTTCGAGAGGAC 4631
2305 GGAGGCACTGACAGTTGCCAGGCTGACAGTGAAGGCTCTGTGTTTCTTCGAGAGGAC 2364
4632 AAATACATTTTACAGAGTCACTTCTGGGCTGTGGCTGTGCACCCCAATAAGCCCT 4691
2365 AAATACATTTTACAGAGTCACTTCTGGGCTGTGGCTGTGCACCCCAATAAGCCCT 2424
4692 GGTGTCAAGCTGTGTTCAAGGTTTGTGTTACTGTGAGTGAAGAGTAAGTAAGTAAT 4751
2425 GGTGTCAAGCTGTGTTCAAGGTTTGTGTTACTGTGAGTGAAGAGTAAGTAAGTAAT 2484
4752 TAAATGGACGGAGACAGAGTGAAGCATCAACTTAGAAGCTGAAGGTGGTAAAG 4811
```

[illegible][illegible]

Query Match	24.7%	Score 1235.4;	DB 6;	Length 2753;
Best Local Similarity	80.9%;	Pred. No. 0;		
Matches 1511; Conservative	0;	Mismatches 276;	Indels 80;	Gaps 3;

Qy	3159	TATCAGGAGCATTCCTCCACACAGTGCACAGGAAAGACCTCTCAATGTGGTCATATAG	3218
Db	911	TATCGGGGGAATGTGGCTGTACCGTGCCTGGGACACCTGTACAGCACTGGATGACAG	970
Qy	3219	ATTACACACNTGSCATCCAGAGAGACCCGAAATCTTACCAAATGCTGGCTGACCCGAGAC	3278
Db	971	ACCCCTCACACATFATCAGGACACCGAAACCTTCCCTGCAAAATTTGGATGTAAGAC	1030
Qy	3279	TACTGCAGATCCAGATTCTGGGAAACAACCCCTGTGTATACACAACGATCCGTGTGTG	3338
Db	1031	TACTGGCCGATCTCTGAGGGAAGAAAGGGCCCATGTGTCCCTTACACCAACCAACCCAAAGG	1090

OY	3339	AGGGGAGGACTGGCAATCTACACCAATGCTGAGAAACAGAAATCAGGTGCTTAGAGACT	33398
Db	1091	CGGGGGAGTACTGTAAATACCTCTCTG-----	1119
OY	3399	CCCACTGTTGTTCCAGTTTCCAGAGATGGAGGCTATTCTGAAAGCAGACCAATGAGCA	34585
Db	1120	-----TGACTCTCTCCCAAGTATCCAGGGAAACAAATTTGGCTCCCGACACCACTGAGCTA	11774
OY	3459	ACCCCTGTGGTCCGGCACTGCTACCAATGTAATGGCCAGAGTTATCGAGGACATATCTCC	3518
Db	1175	ACCCCTGTGGTCCAGAGACTGCTACCAATGGATGGATGGACAGAGCTACCGAGGACATCTCTCC	12344
OY	3519	ACCACTGTACAGAGAGACATGTGCATCTTGGTATCTCATGACACACACCGGCATCAG	35787
Db	1235	ACCAACCAACCAAGAAAGAAAGTGCATCTTGGTATCTGTATGACACCAACCGGGACAG	12944
OY	3579	AGGACCCCAAAAAAATACCCAAATGAATGGCTGACAAATGAATCTCTCCAGAAATCCAGAT	36388
Db	1295	AGGACCCCAAAAAAATACCCAAATGAATGGCTGACAAATGAATCTCTCCAGAAATCCAGAT	13544
OY	3639	GCCGATACAGGCCCTTGGTGTATTTTACCAATGGAACCCACAGATCAAGTGGAGTACTCAAC	36989
Db	1395	GCCGATACAGGCCCTTGGTGTATTTTACCAACAGACCCACAGCTCAGTGGAGTACTCAAC	14144
OY	3699	CTTGACCGGATGCTCGAGACAGGAAGGACATGTGGTGGCTCCCGACGATGTCAACAGT	37588
Db	1415	CTGAAAAAATGCTCAGAAACAGAAAGCAGATGTATGACACTCCGGCTGTGTCTGCTT	14744
OY	3759	CCAAAGCTAGGGCTCTCTTCTGAACAAAGCTGTATGTTGGAAATGGAAAGATACCGG	38188
Db	1475	CCGATGTGTAGAAACATCTTTTCGAAAGAAAGATGTATGTTGGAAATGGAAAGATACCCA	15344
OY	3819	GCCAAAGAGCAACCACTGTTACTGGGAGCCCATGGCAGAGATGGCTGCCAGAGGCC	38787
Db	1535	GCCAAAGAGGCGCACACTGTTACTGGGAGCCCATGCCAGAGCTGGCTGCCAGAGGCC	15944
OY	3879	CATAGACACAGAGACTTATCTCCAGGGCAAAATAAATGGGAGGCTCGAAAAAATTAC	39388
Db	1595	CATAGACACAGACTTATCTCCAGAGACAAATCCAGGGGGGTCTGGAAAAAATTAC	16544
OY	3939	TGCCGTAACCTGATGGTGAATCATATGTTCCCTGGTCTACAAATGAATCCAGAAAA	39988
Db	1655	TGCCGTAACCTGATGGTGAATCATATGTTCCCTGGTCTACAAATGAATCCAGAAAA	17144
OY	3999	CTTTTGAATCTGTGAATACCTCTCTCTGTGATCCTCTTATTTGATTTTGGAGGCT	40588
Db	1715	CTTTTGAATCTGTGAATACCTCTCTCTGTGATCCTCTTATTTGATTTTGGAGGCT	17744
OY	4059	CAGTGGAGCCCAAAATATGTCTCGAAAGCATTTGAAGGGGGGTGTGGGCCACCCACAT	41188
Db	1775	CAGTGGAGCCCAAAATATGTCTCGAAAGGTTTGAAGGGGGGTGTGGGCCACCCACAT	18344
OY	4119	TCTGTGACCTGTGCAGATCAGTCTCAGAACAGGTTTGAAGACCTTGTGGAGGACCC	41787
Db	1835	TCTGTGACCTGTGCAGATCAGTCTTGAACAAAGTTTGAAGACCTTGTGGAGGACCC	18944
OY	4179	TTTATATATCCCAAGATGGGTCTCTACATCGTGCATCCTCTTGAAGAGTCTCTAAGGCT	42388
Db	1895	TTTATATATCCCAAGATGGGTCTCTACATCGTGCATCCTCTTGAAGAGTCTCTAAGGCT	19544
OY	4239	TGATCTTACAAGGTATCTCTGGGTGCAACACAGAAAGTGAACCTGCAATCTCATGTTCAG	42988
Db	1955	TGATCTTACAAGGTATCTCTGGGTGCAACACAGAAAGTGAATCTGCAACCGCATGTTCAG	20144
OY	4299	GAAATAGAAATCTCAGGCTGTTCTTGGAGCCCAACAGACAGATTAATGCTTGTCTTAAAG	43588
Db	2015	GAAATAGAAATCTCAGGCTGTTCTTGGAGCCCAACAGAAAGTATTAATGCTTGTCTTAAAG	20744
OY	4359	CTTAAAGAGGCTCGGCTATCTACTGACAAAGTAATAGCCAGTTCTGCAATCCCGACAGC	44188
Db	2075	CTTAAAGAGGCTCGGCTATCTACTGACAAAGTAATCCAGTTTCTGCAATCCCGAAT	21344
OY	4419	TGATGTGATCCCGCAGAGATGAAATGTTATATCATCTGCTGGGAGAAACCAAGGTAC	44787

```
Db 2135 TATGTGGTGGTGCACGGAGCCGAATGTTTCATCACTGGCTGGGAGAAACCCAGGTAAT 2194
      || ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 4479 TTTGGAGCTGGGCTTCTCAAGAAAGCCAGCTCCCTTTATAGATGAATGAAGTGCAT 4538
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2195 TTTGGAGCTGGGCTTCTCAAGAAAGCCAGCTCCCTTTATAGATGAATGAAGTGCAT 2254
Oy 4539 CACTATAGT-----ATAATGTGCTGACGATTTTGGGC 4571
      | ||||| |||
Db 2255 CCAATCAATTTTCAAGATGGAAGATCCAACTGCTGCTGGCCATTAAGCCT 2314
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 4572 AGAGCACTGACAGTTGGCCAGGTGACAGTGAAGAGGCTCTGGTGTCTTCGAGAAGAC 4631
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2315 GGAGCACTGACAGTTGGCCAGGTGACAGTGAAGAGTCTCTGGTGTCTTCGAGAAGAC 2374
Oy 4632 AATTCATTTTCAAGAGTCACTCTTGGGGTCTTGGCTGACGCCCAATTAAGCCT 4691
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2375 AATTCATTTTCAAGAGTCACTCTTGGGGTCTTGGCTGACGCCCAATTAAGCCT 2434
Oy 4692 GGTGTCTATGCTCGTGTTCAGGTTTGTACTGTGATGAGGGAATGATGAAATTAAT 4751
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2435 GGTGTCTATGCTCGTGTTCAGGTTTGTACTGTGATGAGGGAATGATGAAATTAAT 2494
Oy 4752 TAATTTGAGCGGAGACAGTGAAGCAACCTTAAGATTAAGCAAGTGAAGTGAAG 4811
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2495 TAATTTGAGCGGAGACAGTGAAGCAACCTTAAGATTAAGCAAGTGAAGTGAAG 2554
Oy 4812 ATTTAGCATGCTGGAATATATAGACAGCAATCAACAGAGACAGTTCACCACTACAG 4871
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2555 ATTTAGCATGCTGGAATATATAGACAGCAATCAACAGAGACAGTTCACCACTACAG 2614
Oy 4872 CTATGCAACACCTTGCAATTTTGTGTAATTTTGTATGAAGCTTTAAGGCTGACATGAC 4931
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2615 CTAACGCCAACAACCTGCGCATTTTGTGTGTATTT-----TTTCAATGCT 2657
Oy 4932 AAATCTGATTTAAGCTCTACATGCTATGACATTTGTTAAAAATTAACCTGCACTTAAT 4991
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2658 GGAATCTGTAGTAAGTACATACCTATGACATTTGTTAAAAATTAACCTGTAATTAC 2717
Oy 4992 TTGATTT 4998
      || |||
Db 2718 TTTGATTT 2724
      || |||

RESULT 6
HDPKGM
LOCUS HDPKGM 1784 bp mRNA linear PRI 08-JAN-1995
DEFINITION Human plasminogen mRNA, 3' end.
ACCESSION K02922
VERSION K02922.1 GI:190112
KEYWORDS glycoprotein; plasminogen.
SOURCE Human liver, cdna to mRNA, library of S.L.Woo.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
  1 (bases 1 to 1784)
  Mallinowski,D.P., Sadler,J.E. and Davie,E.W.
  Characterization of a complementary deoxyribonucleic acid coding
  for human and bovine plasminogen
JOURNAL Biochemistry 23 (18), 4243-4250 (1984)
MEDLINE 85023311
PUBMED 6148961
FEATURES
  source
    1..1784
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /map="626-q27"
    1..1784
    /gene="PLG"
    /gene="PLG"
    /gene="PLG"
    /note="pmg mRNA; G00-119-498"
  CDS
    <1..1560
```

```
/gene="PLG"
/note="plasminogen"
/codon_start=1
/protein_id="AAA60124.1"
/db_xref="GI:387031"
/db_xref="GDB:G00-119-498"
/translation="SGHTCOHMSAOCPPHHNRPENPCKNLDENYCNRPDKRAPMC
HTNSQVMEYCKTPSCDSSPEEDLAPAPPELTPVPODCYHGDGQSYRSTSTT
TGKCKQSMSSMTKPHRKHOKTPENPMAGLIMTNCRNPDADKQWCTTTPSTRMEYCNL
KKGSGTEASVAPPPVYLLPDEVETPSEDDMEGNGKIRGKFTATVTGPGCDMAOE
PHRSHIFPETNPRAAGLEKNYCNRPDGGVGGPCTYTNPKLTDVCDVPGQAAFPDC
GKPOVEPKPCGRVVGCVAHPRSMQVSLRTREGMHRCGGTLLSPMVLTAACLE
KSPRSYKVLVIAHDEVLNEPVOETEVSRLEFETREDIALISPAVATTKVIP
ACLSPNVVADRTRECFITGMGETGFGAGLKEQLVILNKVCNRYEELNGVOS
TELCAGHLAGGTDCQGDSSGPIVCFEKDYILIOVTSMLGCLCAPNRPGVYVRSRF
VTWTEGVARN"
BASE COUNT 473 a 452 c 457 g 402 t
ORIGIN
Query Match 24.7%; Score 1232.8; DB 9; Length 1784;
Best Local Similarity 81.8%; Pred. No. 0;
Matches 1494; Conservative 0; Mismatches 252; Indels 80; Gaps 3;

Oy 3195 ACCTGCAATCTTGGTCATCTATGATACCAACATGCAATGAGAGCCAGAAAACCTAC 3254
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10 ACCTGCAATCTTGGTCATCTATGATACCAACATGCAATGAGAGCCAGAAAACCTAC 69
Oy 3255 CCAATGCTGGCTGACCGAGAACTACTGAGAAATCCAGATTCGTGGAAACACCTGG 3314
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 70 CCTGCAAAAATTTGGATGAAACATACGCGCAATCTCTGACGGAAGAGGCCCATGG 129
Oy 3315 TGTTACCAACCGATCCGTTGTGAGTGGGGAGTACTGATCTGACATCTGACAAATGTCAGAA 3374
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130 TGCCATACCAACCAACGCAATGGCGTGGAGTACTGTAAAGATACCGTCTCTG----- 182
Oy 3375 ACAGAATCAGGTCTCTAGAGACTCCACGTGTGTCAGTTCCAGACATGAGAGCTCAT 3434
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 183 -----TGACCTCTCCCAATGATCCAGGAACAAATGG 213
Oy 3435 TCTGAAGCAGACCACTGAGCAAAACCCCTGGTGGCGGAGTGTCTACATGTAATGGC 3494
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 214 GCTCCACAGACACACACTAGTAAACCCCTGGTGGCGGAGTGTCTACATGTAATGGC 273
Oy 3495 CAGAGTTATGAGGCACTTCTCCACACTGTACAGAGAAAGATATGTCATCTTGGTCA 3554
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 274 CAGAGTTATGAGGCACTTCTCCACACTGTACAGAGAAAGATATGTCATCTTGGTCA 333
Oy 3555 TCCATGACACACACCGGATGAGAGAACCCAGAAAACCTACCAATGATGAGCTGTACA 3614
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 334 TCTATGACACACACCGGATGAGAGAACCCAGAAAACCTACCAATGATGAGCTGTACA 393
Oy 3615 ATGAATCTACTGCAAGATATCAATGATGAGAGCTGTGTTTATTCATGAGACCC 3674
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 394 ATGAATCTACTGCAAGATATCAATGATGAGAGCTGTGTTTATTCATGAGACCC 453
Oy 3675 AGCATAGGTGGGAGTACTGCAACCTGACGAGATCTCAGACAGAGAGGAGTGTGTC 3734
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 454 AGGTACAGGTGGGAGTACTGCAACCTGAGAAAATGCTCAGGAACAGAGAGAGTGTGA 513
Oy 3735 GCTCTCCGACTGTCAATCAGAGTTCGAAGCTAGAGGCTCTCTTGTGAACAAGACTGTATG 3794
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 514 GCACCTCCGCTGTCTGCTGCTCCAGATGTAAGATCTTCCGAAAGAGACTGTATG 573
Oy 3795 TTTGGAAATGGGAAAGATACCGGGGCAAGAGCAACATGTTACTGAGAGCCATGC 3854
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 574 TTTGGAAATGGGAAAGATACCGAGGCAAGAGGAGCAATGTTACTGAGAGCCATGC 633
Oy 3855 CAGGAATGGGCTCCCAAGAGCCCATAGACACAGCTTCAATCCAGGAGCAAAATATA 3914
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 634 CAGGAATGGGCTCCCAAGAGCCCATAGACACAGCTTCAATCCAGGAGCAAAATATA 693
Oy 3915 TGGCAGGTGTGGAATAAATTAATGCGTAACCTGATGATGATCAATCAATGTCCTGG 3974
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 694 CGGGGGGCTGAAAAAATTAACCTGTAACCTGATGATAGTGTGCTCCG 753
Oy 3975 TGTACAAATGAATCCAAAAAATTTTGTACTGATGATATCCCTCTGTGCAATC 4034
Db 754 TGTACACAGCAAAATCCAAAAAATTTTGTACTGATGATATCCCTCTGTGCAATC 813
Oy 4035 TCTTCATTTGATTTGGGAAGCTCAAGTGGAGCCGAGAAATGTCTGGAAGATTTGTA 4094
Db 814 CCTTCATTTGATTTGGGAAGCTCAAGTGGAGCCGAGAAATGTCTGGAAGATTTGTA 873
Oy 4095 GGGGGGTGTGGGGCCCACTTCCTGCGCTGCAAGTGTGATGCAAGAAAGTTT 4154
Db 874 GGGGGGTGTGGGGCCCACTTCCTGCGCTGCAAGTGTGATGCAAGAAAGTTT 933
Oy 4155 GGAAGACATTTGTGTGAGGACCTTAATATCCAGAGAGTGGCTGTGATGCTGTGTA 4214
Db 934 GGAATGACATTTGTGTGAGGACCTTAATATCCAGAGAGTGGCTGTGATGCTGTGTA 993
Oy 4215 TGGTTGAAGAAATCTCAAGGCTTATCTTACAAAGTATCTCTGGGTGACACCAAGAA 4274
Db 994 TGGTTGAAGAAATCTCAAGGCTTATCTTACAAAGTATCTCTGGGTGACACCAAGAA 1053
Oy 4275 GTGAACCTCGAATCTCATGTTTCAGAAATAGAAGTGTCTAGGCTGTCTTGGAGCCACA 4334
Db 1054 GTGAATCTCGAATCTCATGTTTCAGAAATAGAAGTGTCTAGGCTGTCTTGGAGCCACA 1113
Oy 4335 CAACGACATATATGCTTGTCTTAAGTCAAGACAGCCCTGCGGTATCTACTGACAAAGTATG 4394
Db 1114 CGAAAGATATATGCTTGTCTTAAGTCAAGACAGCCCTGCGGTATCTACTGACAAAGTATG 1173
Oy 4395 CCAAGCTTGTCTGCAATCCCAAGACATGATGTCACGCGCAGATGATGTTTACATCT 4454
Db 1174 CCAAGCTTGTCTGCAATCCCAAGACATGATGTCACGCGCAGATGATGTTTACATCT 1233
Oy 4455 GGGGGGGAAGAAACCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4514
Db 1234 GGGGGGGAAGAAACCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1293
Oy 4515 GTTATGAGATGAAGTGTGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 4548
Db 1294 GTTATGAGATGAAGTGTGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1353
Oy 4549 -ATATTTGTGCTGAGATTTGGCCAGAGGACATGACAGTGTGACAGGAGAGAGGAGG 4607
Db 1354 GAACTGTGTCTGTGGCAATTTGGCCAGAGGACATGACAGTGTGACAGGAGAGAGGAGG 1413
Oy 4608 CTTCTGTGTGTGTGCAAGAGCAAAATCAATTTTACAAGAGTCACTTCTTGGGCTCT 4667
Db 1414 CTTCTGTGTGTGTGCAAGAGCAAAATCAATTTTACAAGAGTCACTTCTTGGGCTCT 1473
Oy 4668 GGGTGTGCTGAGCCCAATAGCTGTGTCTATGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 4727
Db 1474 GGGTGTGCTGAGCCCAATAGCTGTGTCTATGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 1533
Oy 4728 ATTGAGGGAATGATGAGAAATTAATTTGAGCGGAGACAGAGTGAAGATCAACCTTAC 4787
Db 1534 ATTGAGGGAATGATGAGAAATTAATTTGAGCGGAGACAGAGTGAAGATCAACCTTAC 1593
Oy 4788 TTGAAAGCTGAAGCTGTGGATTTAGATGCTGTGGAATTAATTAAGACATCAATCAAC 4847
Db 1594 CTGAGAGCTGGAAGAGGATAGGATTTAGATGCTGTGGAATTAATTAAGACATCAATCAAC 1653
Oy 4848 GAAGACATGTTCCCAAGTACAGAGTATGCCAAACCTTGGCAATTTTGTGTGTGTGTGT 4907
Db 1654 GAAGACATGTTCCCAAGTACAGAGTATGCCAAACCTTGGCAATTTTGTGTGTGTGTGTGT 1709
Oy 4908 ATAAAGCTTTAAAGTGTGACAAATTTCTATTAATTAATTAATTAATTAATTAATTAAT 4967
Db 1710 ATAAAGCTTTAAAGTGTGACAAATTTCTATTAATTAATTAATTAATTAATTAATTAAT 1756
Oy 4968 TTAATAATTAATTAATTAATTAATTTT 4993
Db 1757 TTAATAATTAATTAATTAATTAATTT 1782

RESULT 7
106212
LOCUS 106212 2679 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 14 from Patent EP 031944.
ACCESSION 106212
VERSION 106212.1 GI:590227
KEYWORDS
SOURCE
ORGANISM unknown.
REFERENCE
AUTHORS 1 (bases 1 to 2679)
Mullvihill,E.R., Berkner,K.L., Foster,D.C., Kumar,A.A., Mackay,V.L.
and Parker,G.E.
TITLE Co-expression in eukaryotic cells
JOURNAL Patent: EP 031944-A2 14 14-JUN-1989;
FEATURES
source Location/Qualifiers
BASE COUNT 750 a 664 c 670 g 595 t
ORIGIN
Query Match 24.6%; Score 1228.8; DB 6; Length 2679;
Best Local Similarity 80.8%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 277; Indels 80; Gaps 3;
Oy 3159 TATCGAGGCATATCTCCACCATCTGTCACAGAGAGACCTGTCAATCTTGTCTATATG 3218
Db 868 TATCGAGGCATATCTCCACCATCTGTCACAGAGAGACCTGTCAATCTTGTCTATATG 927
Oy 3219 ATACACACCTGTCATCAGAGAGACCCAGAAATCTACCAAAATGCTGCTGACGAGAAC 3278
Db 928 ACCCTTCACACATTAACAGAGACACGAAATCTTCCCTCGCAAAATTTGATATAAAAC 987
Oy 3279 TACTCGAGGAATCCAGATTTCTGGGAAACAACCTGTGTGTATACAAACGATCCGTGTG 3338
Db 988 TACTCGAGGAATCCAGATTTCTGGGAAACAACCTGTGTGTATACAAACGATCCGTGTG 1047
Oy 3339 AGTGGAGTACTGATCTGACATCTGACATCTGACAAACAGAAATCAAGGTGTCTAGAGCT 3398
Db 1048 CGGTGGAGTACTGATCTGACATCTGACATCTGACAAACAGAAATCAAGGTGTCTAGAGCT 1076
Oy 3399 CCCACTGTTGTTCCAGTTCAGAGCAGAGGCTCATCTTGAAGACACCAATGAGCA 3458
Db 1077 -----TGACTCCTCCAGATATCCAGGAACAATTTGGCTCCACAGCACCACTGAGCTA 1131
Oy 3459 ACCCTGTGTGTCGCGAGTGTACATGTAATGGCCAGAGTATATGAGGACATTTCTCC 3518
Db 1132 ACCCTGTGTGTCGCGAGTGTACATGTAATGGCCAGAGTATATGAGGACATTTCTCC 1191
Oy 3519 ACCACTGTACAGAGAGACATGTCAATCTTGTGTCATCCATGACACACACCCGAGTACG 3578
Db 1192 ACCACTGTACAGAGAGAGACATGTCAATCTTGTGTCATCCATGACACACACCCGAGTACG 1251
Oy 3579 AGGACCCAGAAACTACCCAAATGATGGCTGTACAAATTAATTAATTAATTAATTAATTAAT 3638
Db 1252 AAGACCCAGAAACTACCCAAATGATGGCTGTACAAATTAATTAATTAATTAATTAATTAAT 1311
Oy 3639 GCCGATACAGGCTTGTGTGTTTACATGAGACCCAGATAGTGGAGTACTGCAAC 3698
Db 1312 GCCGATACAGGCTTGTGTGTTTACATGAGACCCAGATAGTGGAGTACTGCAAC 1371
Oy 3699 CTGACGAGATCTGACAGACAGAGAGAGTGTGTGCTCCAGACTGATCAAGATT 3758
Db 1372 CTGAAAAATGCTGAGGAAGAGAGAGTGTGTGACACCTCCGCTGTGTCTGTGT 1431
Oy 3759 CCAAGCTTAGGCTCTTCTTGAACAAGATGTATTTTGGATGGAAGAGATATCCG 3818
Db 1432 CCAGATGTAGAGACTCTTCCGAAAGACTGTATTTTGGAAATGGGAAAGGATATCCGA 1491
Oy 3819 GGCAAGAGGCATCTGTTACTGGGAGCCCTGACAGAGTGGAGTGGCCAGAGGCC 3878
Db 1491 GGCAAGAGGCATCTGTTACTGGGAGCCCTGACAGAGTGGAGTGGCCAGAGGCC 3878

Db 960 ACCCTCAGACATTAACAGACACCAAAAATTTCCTGCAAAAATTTGGATGAAAC 1019
Oy 3279 TACTGACAGATTCAGATTCCTGGGAAACAACCTGGTGTATACACAGCATCCGGTGTG 3338
Db 1020 TACTGCGCAATCTGTATGAGAAAGGCCCATGGTGTGTATACAAACCAACGCAAGT 1079
Oy 3339 AGTGGAGATCTGCAATCTGACACAATGCTGAGAAAAGAAATGAGTGTCTAGAGACT 3398
Db 1080 CGGTGGAGATCTGTAGATACCGTCTG----- 1108
Oy 3399 CCGATGTGTTCAGTCCAGACATGAGAGCTCATTTCTGAGCAGACCACTGAGCAA 3458
Db 1109 -----TGAGTCTCCACAGATTCACGGAACCATTTGGATCCACAGCCATGAGTTT 1163
Oy 3459 ACCCTGTCGTCGCGAGTGTCTACATGTAATGGCAGAGATTATCAGAGCATTTCTCC 3518
Db 1164 ACTCTGTGGTCCAGAGTGTCTACATGGTGTGAGGAGGAGCTACCGAGGACATCTCTCC 1223
Oy 3519 ACCACTGTCAAGAGAGACATGTCAATCTGTGTCATCCATACACACACCGGCATTCAG 3578
Db 1224 ACCACACACAGAGAAAGATGTCACTGTGTCATCTATGACACACACTGGCATGAG 1283
Oy 3579 AGGACCCAGAAAACCTCCAAATGCTGGCTGACATGACTACTGAGAGAAATCCAGAT 3638
Db 1284 AAGACCCAGAAAACCTCCAAATGCTGGCTGACATGACTACTGAGAGAAATCCAGAT 1343
Oy 3639 GCCCATAGAGCCCTGTGGTGTTCATCATGAGACCCAGCATCAGGTGAGTACTGCAAC 3698
Db 1344 GCCCATAGAGCCCTGTGGTGTTCATCATGAGACCCAGCATCAGGTGAGTACTGCAAC 1403
Oy 3699 CTGACCGATGCTCAGACAGAGAGACTGTGTGCTGCTCCAGACTGTATCCAGATT 3758
Db 1404 CTGAAAAATGCTCAGAGAAAGAGAGTGTGAGCAGCTCCGCTGTGGCCCACTT 1463
Oy 3759 CCAAGCCATGAGGCTCCTCTGTGAACAAGACTGTATGTTGGGAAAGGGAATGATCCG 3818
Db 1464 CCAATGACAGACTCTCTCCAGAGAACTGTATGTTGGGAAAGGGAATGATCCG 1533
Oy 3819 GGCATAGAGCAACACTGTACTGTGAGCCAGCCAGCAGAAATGGGTCGCCAGAGCCC 3878
Db 1534 GGCATAGAGCAACACTGTACTGTGAGCCAGCAGCAGAAATGGGTCGCCAGAGCCC 1583
Oy 3879 CATGACAGACAGACTTCCAGGGCAAAATTAATGGGAGGTCTGGAATAAATTTAC 3938
Db 1584 CACAGCCACCGCATTTCACTCCAGAGCAAAATCCACGGGAGGTCTGGAATAAATTTAC 1643
Oy 3939 TCCGCTAACCTGATGTGATCAATGCTCCCTGCTGTACATATGATCCAGAAA 3998
Db 1644 TCCGCTAACCTGATGTGATGAGGTGCTCCCTGCTGTACATCAAGAAA 1703
Oy 3999 CTTTTCGACTACTGTATCCCTCTCTGTGATCCTCTTCATTTGATTTGGAGACCT 4058
Db 1704 CTTTTCGACTACTGTATGTCCTCTGAGTGTGGGCTCTTCATTTGATTTGGAGACCT 1763
Oy 4059 CAATGAGCGGAAGAAATCTCTGAGAGCATTTGAGGGGTGTGTGGCCACCCACAT 4118
Db 1764 CAATGAGCGGAAGAAATCTCTGAGAGGTGTAAGGGGTGTGTGGCCACCCACAT 1823
Oy 4119 TCTGAGCCCTGAGCAAGTCTCAGACAGAGATTTGAGAAACACTTCTGTGAGGACCC 4178
Db 1824 TCTGAGCCCTGAGCAAACTCTTAAACAGAGCTTGGAAATCACTTCTGTGAGGACCC 1883
Oy 4179 TTAAATATCCAGAGTGGGTGATGCTGTGTCATCTGTGAAAGAAATCCCTCAAGGCT 4238
Db 1884 TTGAATATCCAGAGTGGGTGATGCTGTGTCATCTGTGAAAGAAATCCCTCAAGGCT 1943
Oy 4239 TCATCTCAAGAGTCTCTGGTGCACACAAAGATGAACTTCATCTCATGTTTCAG 4298
Db 1944 TCATCTCAAGAGTCTCTGGTGCACACCAAGATGAACTTCATCTCATGTTTCAG 2003
Oy 4299 GAATATGAGTGTAGGCTGTCTGTGAGCCACAAAGAGATATGCTTCTCTAAG 4358
|||||

Db 2004 GAATATGAGTATCTAAGATGTTCTCGAGACCCCGAAGAGAGATATTCCTGTAAG 2063
Oy 4359 CTAAGACGCTCGCCGATCATCTACACAAAGTAAATGACAGTGTCTGCTCCACAGAC 4418
Db 2064 CTAAGACGCTCGCCGATCATCTACACAAAGTAAATGACAGTGTCTGCTCCACAGAC 2123
Oy 4419 TACATGCTCACCGCAGAGACTGAATGTTATCATCTGCTGGGAGAAACCAAGTATCC 4478
Db 2124 TATGTGCTCGCTGACCGAGCAATGTTATCATCTGCTGGGAGAAACCAAGTATCC 2183
Oy 4479 TTTGGAGTGGCTCTTCAAGAAAGCCAGCTCTGTTATTAAGAAATGAATGTCAT 4538
Db 2184 TATGGGCTGGCTCTTCAAGAAAGCCGCTCCCGTATTAAGAAATGAATGTCAT 2243
Oy 4539 CACTATAGT-----ATATTTGGCTGAGCATTTGGCC 4571
Db 2244 CGTATGAGTTCGTAATGAGACATCAAAACCCAGCTGTGTGGGAGCATTTGGCC 2303
Oy 4572 AAGACATGACATGTCAGAGGTGACAGTGAAGGAGGCTGTGTTGCTTGAGAGAGAC 4631
Db 2304 GGAGGACATGACATGTCAGAGGTGACAGTGAAGGAGGCTGTGTTGCTTGAGAGAGAC 2363
Oy 4632 AAATACATTTTAAAGAGTCACTCTTGGGCTGTGGCTGAGACGCCCAATTAAGCT 4691
Db 2364 AAATACATTTTAAAGAGTCACTCTTGGGCTGTGGCTGAGACGCCCAATTAAGCT 2423
Oy 4692 GGTGTCTATGCTGTGTTCAAGGTTGTTACTGTGATTAAGGAAATGATGAAATTAAT 4751
Db 2424 GGTGTCTATGCTGTGTTCAAGGTTGTTACTGTGATTAAGGAAATGATGAAATTAAT 2483
Oy 4752 TAAATGACAGGGA-GACAGAGTAAACATCAACTTAAAGCTGAAGCTGGGTAAG 4810
Db 2484 TAAATGACAGGGAATACAGAGTAAACATCAACTTAAAGCTGGGTAAG 2543
Oy 4811 GATTAGCATGCTGGAATTAATGACAGAAATCAAGCAACACTGTCCAGACTTACCA 4870
Db 2544 GATTAGCATGCTGGAATTAATGACAGAAATCAAGCAACACTGTCCAGACTTACCA 2603
Oy 4871 GCTATGCCAAACCTTGGCATTTTGGTATTTGTGTATTAAGCTTTAAGCTGACTGA 4930
Db 2604 GGGAAAGCAAACTTACAGCATTTTGTATTAAT-----TTTGACTGTC 2646
Oy 4931 CAAATCTGATTAAGGTGATAGCTATGA-CATTTGTTAAATAAATCTGCACTTA 4989
Db 2647 TGAATTTGTAATTAAGGTGATAGCTATGACCATTTTGTAAATAAATCTGCTACTTA 2706

RESULT 9
ARI05749 2296 bp DNA linear PAT 14-FEB-2001
LOCUS ARI05749
DEFINITION Sequence 18 from patent US 6103244.
ACCESSION ARI05749
VERSION ARI05749.1 GI:12819814
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2296)
AUTHORS Dörner,F., Scheifflinger,F., Falkner,F., Gunter, and Pfeleiderer,M.
TITLE Methods for generating immune responses employing modified vaccinia
of fowlpox viruses
JOURNAL Patent: US 6103244-A 18 15-AUG-2000;
FEATURES
location/Qualifiers
1..2296
BASE COUNT 622 a 597 c 583 g 494 t
ORIGIN
Query Match 22.2%; Score 1110.6; DB 6; Length 2296;
Best Local Similarity 81.4%; Pred. NO. 0;
Matches 1340; Conservative 0; Mismatches 244; Indels 63; Gaps 2;
Oy 3159 TATCGAGCATATCTCCACACTGTGCACAGAGAGACTGTCAATCTGTGATCATATG 3218


```

Db 636 TATCGGGGAAATGTGGCTTTACGTTTCGGGCAACCTGTGACGACTGGATGSCACG 695
Qy 3219 ATACCACTGGCATCAGAGACCCAGAAAACATACCCAAATGCTGGCTCGACGAGAC 3278
Db 696 ACCCTCAGACACATTAACAGGACACCAAAAACCTCCCTCCAAAATTTGGATGAAAAAC 755
Qy 3279 TACTGAGGAATTCAGATTCTGGAAACACCTGGGTTTACACACGATCCGTGCTG 3338
Db 756 TACTGCGCAATCTGACGAAAAAGGCCCCCATGCTGCTACACCAACAGCAAGTG 815
Qy 3339 AGGTGGAGTACTGCAATCTGACACATGCTGAGAAACAGATCGTCTGAGACAT 3398
Db 816 CGGTGGAGTACTGTAAGTATACCTGCTCCTG----- 844
Qy 3399 CCCAGTGTGTTCCAGATTCCAGCATGGAGGCTCATTCGAAAGCAGCACCAGTGCACAA 3458
Db 845 -----TGACTCCTCCCAAGTATCCAGGACAAATTTGGCTCCACAGCACCAGTGCAC 899
Qy 3459 ACCCTGTGGTCCGGCAGTGTACCATGCTAATGGCCAGTAAWCGAGCACAATTCCTCC 3518
Db 900 ACCCTGTGGTCCAGAGTGTACCAACGATGAGACAGTACCGAGCAGACATCCTCC 959
Qy 3519 ACCAGTGCAGAGAGACATGTCATCTGGTATCCTCATCAGACACACACCGGACATGAC 3578
Db 960 ACCACACACACAGAGAAAGATGTGATGCTGTGATGTATGACACACACACCGGACACG 1019
Qy 3579 AGAGACCCAGAAAACTACCCAAATGATGGCTGACATGAATCTAGTCAGAAATCCAGAT 3638
Db 1020 AAGACCCAGAAAACTACCCAAATGATGGCTGACATGAATCTAGTCAGAAATCCAGAT 1079
Qy 3639 GCGGATACAGGCCCCCTGGTGTTCACATGGACCCACATCAGATGAGAGTACTGCAAC 3698
Db 1080 GCGGATACAGGCCCCCTGGTGTTCACATGGACCCACATCAGATGAGAGTACTGCAAC 1139
Qy 3699 CTGACGGAGTGTGACACACAGAGAGTGTGCTGCTCCGATGTCATCCAGGTT 3758
Db 1140 CTGAAAAAATGCTCAGAGAACAGAGAGTGTGATGACCTCCGCTGTGCTCCTCTT 1199
Qy 3759 CCAAGCTAGGGGCTCTCTCTGACACAGACTATGTTGGAAATGGAAAGATACCG 3818
Db 1200 CCAAGATGTAGAGACTCTCTCCGAGAGAGACTGTATGTTGGAAATGGAAAGATACCGA 1259
Qy 3819 GCGAAGAGGCAACCACTGTACTGAGAGCCATGCCAGATGGGCTGCCAGAGACC 3878
Db 1260 GCGAAGAGGCGACCACTGTACTGAGAGCCATGCCAGACTGGGCTGCCAGAGACC 1319
Qy 3879 CATGACACAGCAGCTTCATTCAGGAGACAAATTAATGGGAGTCTGSAAAAAATTAC 3938
Db 1320 CATGACACAGCATTTCATTCAGAGACAAATTCACGCGGGGCTGSAAAAAATTAC 1379
Qy 3939 TGGCGTAACTGATGATGATCAATGATGCTGCTGCTACACAAATGATCCAGAAA 3998
Db 1380 TGGCGTAACTGATGATGATGATGATGCTGCTGCTACACAAATGATCCAGAAA 1439
Qy 3999 CTATTGACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4058
Db 1440 CTATTGACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1499
Qy 4059 CAAGTGAAGCCGAAAGATGCTGAGAGCAATGTAAGGGGGTGTGGGCCACCCACAT 4118
Db 1500 CAAGTGAAGCCGAAAGATGCTGAGAGGGTGTGGGGGGTGTGTGGGCCACCCACAT 1559
Qy 4119 TCCGAGCCCTGGCAAGTCTCAGAACAGTGTGAAAGACCTTCTGAGAGGACC 4178
Db 1560 TCCGAGCCCTGGCAAGTCTTGAACAAAGTGTGAAAGTGTGAGAGGACC 1619
Qy 4179 TTAATATCCCAAGTGGTGTGATGCTGCTGCTGCTGCTGAGAGTCTCAAGGCT 4238
Db 1620 TTAATATCCCAAGTGGTGTGATGCTGCTGCTGCTGCTGAGAGTCTCAAGGCT 1679
Qy 4239 TCAATCTCAAGTGTCAATCTGGTGCACACCAAGAGTGAATCTGATCTCATGTTGAG 4298

```

```

Db 1680 TCAATCTCAAGTGTCAATCTGGTGCACACCAAGAGTGAATCTGAAACCGATGTGAG 1739
Qy 4299 GAAATGAGATGTCTAGGCTGTTCTTTGGAGCCACACACAGAGATATTTGCTGCTAAG 4358
Db 1740 GAAATGAGATGTCTAGGCTGTTCTTTGGAGCCACACAGAGAGATATTTGCTGCTAAG 1799
Qy 4359 CTAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4418
Db 1800 CTAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1859
Qy 4419 TACTGTGACCCGACAGTGAATGTTACATCAGTGGCTGGGAGAAACCAAGTACC 4478
Db 1860 TACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1919
Qy 4479 TTTGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4538
Db 1920 TTTGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1979
Qy 4539 CACTATAGT-----ATATTTGCTGACCATTTGGCC 4571
Db 1980 CGTATGAGTGTCTGAAATGGAAGATCCATCCAGCAACTCTGCTGGGCAATTTGGCC 2039
Qy 4572 AGAGCAGTACAGATTGCCAGGCTGACAGTGAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 4631
Db 2040 AGAGCAGTACAGATTGCCAGGCTGACAGTGAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2099
Qy 4632 AAATACATTTTACAGAGTCACTCTTGTGGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 4691
Db 2100 AAATACATTTTACAGAGTCACTCTTGTGGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 2159
Qy 4692 GGTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4751
Db 2160 GGTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2219
Qy 4752 TAATTGGACGGAGACAGAGTGAACA 4778
Db 2220 TAATTGGACGGAGACAGAGTGAACA 2246

RESULT 10
LOCUS 114104 2296 bp DNA Linear PAT 26-SEP-1995
DEFINITION Sequence 18 from patent US 5445953.
ACCESSION 114104
VERSION 114104.1 GI:996527
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 2296)
AUTHORS
Dorner, F., Scheifflinger, F. and Falkner, F.G.
TITLE
Direct molecular cloning of a modified poxvirus genome
JOURNAL
Patent: US 5445953-A 18-29-AUG-1995;
FEATURES
Location/Qualifiers
1..2296
BASE COUNT 622 a 597 c 583 g 494 t
ORIGIN
Query Match 22.2%; Score 1110.6; DB 6; Length 2296;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 1340; Conservative 0; Mismatches 244; Indels 63; Gaps 2;
Qy 3159 TATGAGGCAATATCTCCACACTGTCACAGAGAGGAGCTGTCAATCTTGATCATATG 3218
Db 636 TATCGGGGAAATGTGGCTTTACGTTTCGGGCAACCTGTGACGACTGGATGSCACG 695
Qy 3219 ATACCACTGGCATCAGAGACCCAGAAAACATACCCAAATGCTGGCTCGACGAGAC 3278
Db 696 ACCCTCAGACACATTAACAGGACACCAAAAACCTCCCTCCAAAATTTGGATGAAAAAC 755
Qy 3279 TACTGAGGAATTCAGATTCTGGAAACACCTGGGTTTACACACGATCCGTGCTG 3338
Db 756 TACTGCGCAATCTGACGAAAAAGGCCCCCATGCTGCTACACCAACAGCAAGTG 815
Qy 3339 AGGTGGAGTACTGCAATCTGACACATGCTGAGAAACAGATCGTCTGAGACAT 3398
Db 816 CGGTGGAGTACTGTAAGTATACCTGCTCCTG----- 844
Qy 3399 CCCAGTGTGTTCCAGATTCCAGCATGGAGGCTCATTCGAAAGCAGCACCAGTGCACAA 3458
Db 845 -----TGACTCCTCCCAAGTATCCAGGACAAATTTGGCTCCACAGCACCAGTGCAC 899
Qy 3459 ACCCTGTGGTCCGGCAGTGTACCATGCTAATGGCCAGTAAWCGAGCACAATTCCTCC 3518
Db 900 ACCCTGTGGTCCAGAGTGTACCAACGATGAGACAGTACCGAGCAGACATCCTCC 959
Qy 3519 ACCAGTGCAGAGAGACATGTCATCTGGTATCCTCATCAGACACACACCGGACATGAC 3578
Db 960 ACCACACACACAGAGAAAGATGTGATGCTGTGATGTATGACACACACACCGGACACG 1019
Qy 3579 AGAGACCCAGAAAACTACCCAAATGATGGCTGACATGAATCTAGTCAGAAATCCAGAT 3638
Db 1020 AAGACCCAGAAAACTACCCAAATGATGGCTGACATGAATCTAGTCAGAAATCCAGAT 1079
Qy 3639 GCGGATACAGGCCCCCTGGTGTTCACATGGACCCACATCAGATGAGAGTACTGCAAC 3698
Db 1080 GCGGATACAGGCCCCCTGGTGTTCACATGGACCCACATCAGATGAGAGTACTGCAAC 1139
Qy 3699 CTGACGGAGTGTGACACACAGAGAGTGTGCTGCTCCGATGTCATCCAGGTT 3758
Db 1140 CTGAAAAAATGCTCAGAGAACAGAGAGTGTGATGACCTCCGCTGTGCTCCTCTT 1199
Qy 3759 CCAAGCTAGGGGCTCTCTCTGACACAGACTATGTTGGAAATGGAAAGATACCG 3818
Db 1200 CCAAGATGTAGAGACTCTCTCCGAGAGAGACTGTATGTTGGAAATGGAAAGATACCGA 1259
Qy 3819 GCGAAGAGGCAACCACTGTACTGAGAGCCATGCCAGATGGGCTGCCAGAGACC 3878
Db 1260 GCGAAGAGGCGACCACTGTACTGAGAGCCATGCCAGACTGGGCTGCCAGAGACC 1319
Qy 3879 CATGACACAGCAGCTTCATTCAGGAGACAAATTAATGGGAGTCTGSAAAAAATTAC 3938
Db 1320 CATGACACAGCATTTCATTCAGAGACAAATTCACGCGGGGCTGSAAAAAATTAC 1379
Qy 3939 TGGCGTAACTGATGATGATCAATGATGCTGCTGCTACACAAATGATCCAGAAA 3998
Db 1380 TGGCGTAACTGATGATGATGATGATGCTGCTGCTACACAAATGATCCAGAAA 1439
Qy 3999 CTATTGACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4058
Db 1440 CTATTGACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1499
Qy 4059 CAAGTGAAGCCGAAAGATGCTGAGAGCAATGTAAGGGGGTGTGGGCCACCCACAT 4118
Db 1500 CAAGTGAAGCCGAAAGATGCTGAGAGGGTGTGGGGGGTGTGTGGGCCACCCACAT 1559
Qy 4119 TCCGAGCCCTGGCAAGTCTCAGAACAGTGTGAAAGACCTTCTGAGAGGACC 4178
Db 1560 TCCGAGCCCTGGCAAGTCTTGAACAAAGTGTGAAAGTGTGAGAGGACC 1619
Qy 4179 TTAATATCCCAAGTGGTGTGATGCTGCTGCTGCTGCTGAGAGTCTCAAGGCT 4238
Db 1620 TTAATATCCCAAGTGGTGTGATGCTGCTGCTGCTGCTGAGAGTCTCAAGGCT 1679
Qy 4239 TCAATCTCAAGTGTCAATCTGGTGCACACCAAGAGTGAATCTGATCTCATGTTGAG 4298

```

D	756	TATGCGCGAATCTCGACGGAAAAAAGGGCCCATGTGCTGCATTAACCAACAGCCAAATG	815
O	3339	AGGTGGGAGTACTGCATCTGCACAACTGCTAGAAACAGATAGGTGTCTTAGAGACT	3398
D	816	CGGTGGGAGTACTGTAAGATACGGTCTG-----	844
O	3399	CCCACTGTGTTGCCAGTTCACAGATGAGAGCTCATTTGTGAAGCAGCCAACTGAGCA	3458
D	845	-----TGACTCTCTCCCAGTATCCAGGAACAATTGGCTCCCAAGCACCACTAGAGTA	899
O	3459	ACCCCTGTGTCCGGCAGATGTAACATGTAAATGGCCAGATATTCAGAGCAATCTTCC	3518
D	900	ACCCCTGTGTCCAGAGTGTGTACCAAGGTGATGAGACAGACTCCAGAGCAATCTTCC	959
O	3519	ACCACTGTACAGGAAGAGCAATGTCAATCTGTGATCATCAATAGACACACACCGCATCAG	3578
D	960	ACCACACACACAGAAATAAGTGTCAATGTGTGATCATAGACACACACCGCACAG	1019
O	3579	AGCAACCCAGAAATATACCCAATGATGAGTGTGATGACATGAACTGTCCAGAAATCCAGT	3638
D	1020	AABRCCCAAGAAATACCTACCCAATGCTGGCTGTACAAATGAACTCTCAGAGATCCAGAT	1079
O	3639	GCCGATACAGGCCCTTGTGTTTACCATGAGACCCACAGCATCAGGTGGAGTACAGCAAC	3698
D	1080	GCGGATTAAGGCCCTGTGTGTTTACCAACAGACCCACAGTCAAGTGGGAGTACTTCAAC	1139
O	3699	CTGACCCGATGTCTGACACAGAAAGGACGTGAGTGGCTGTCCGATGTGCATCCAGGT	3758
D	1140	CTGAAAAAATGCTCAGAAACAAAGGAGATGTGTATGACACTCCGCTGTGTCTGCTGCT	1199
O	3759	CGAAGCTTAGGCCCTCTTCTGAAACAAGTGTATGTTGGAAATGGAAAGATACGG	3818
D	1200	CCGAGATGTAGAACCTCTCTCGAAGAAAGCTGTATGTTGGAAATGGAAAGATACCA	1259
O	3819	GCGAACAAGCAACACTGTTACTGAGGAGCCCATCCAGAAATGGGTGCCAGAGGCC	3878
D	1260	GCGAAGAGGGCACCACTGTACTGAGGAGCCATCCAGAGTGGGTGCCAGAGGCC	1319
O	3879	CATAGCACAGACGTTTATTTCAGAGSACAAATMAATGGCAGGTCTGGAAAAAATTAC	3938
D	1320	CATAGCACAGATTTTATTACTCCAGAGACAAATCCAGAGGGGGGTCTGGAAAAAATTAC	1379
O	3939	TGCCGTAAACCTGATGTGATCATATGTGTCCCTGTCTCAACAATGAATCCAGAAATA	3998
D	1380	TGCCGTAAACCTGATGTGATGTAGTGTGGTGGTCCGTGTCTCAACAATGAATGAATA	1439
O	3999	CTTTTGTACTATGTGATATCCCTCTCTGTGATATCTCTCATTTGATTTGTGGGAAGCT	4058
D	1440	CTTTACGACTGTGATGTATCCCTCACTGTGTGGGCCCTTATTTGATTTGTGGGAAGCT	1499
O	4059	CAAGTGGAGCCGAAATAATGTCTGGAACATTTGTAGGGGGGTGTGAGCCACCCACAT	4118
D	1500	CAAGTGGAGCCGAAATAATGTCTGGAAGAGTTGTGAGGGGGGTGTGAGCCACCCACAT	1559
O	4119	TTCGTGGCCCTGTGCAAGTCAAGTCAAGCAACAAAGTTTGAAGAAGCTCTGTGGAGCAC	4178
D	1560	TTCGTGGCCCTGTGCAAGTCAAGTCAAGCAACAAAGTTTGAAGAAGCTCTGTGTGGAGCAC	1619
O	4179	TTAATATCCCAAGTGGGTGCTGACTCTGTGCTCACTCTTGAAGAAGTCTCTAAAGGCT	4238
D	1620	TTGATATTCCTCCAGATGGGTGTGACTGTGCTGCCCACTCTTGAAGAAGTCTCTAAAGGCT	1679
O	4239	TGATCTTAACAAGTCACTGTGGGTGACACCAAGAAAGTGAACCTCGAATCTCATTTCTGAC	4298
D	1680	TGATCTTAACAAGTCACTGTGGGTGACACCAAGAAAGTGAACCTCGAAGCTGTGACG	1739
O	4299	GAATATGAAGTCTGAGGCTGTCTTGTGAGCCACACAAAGCAGATATTTGCTTGCCTTAAG	4358
D	1740	GAATATGAAGTCTGAGGCTGTCTTGTGAGCCACCAAGAAAGTATTTGCTTGCCTTAAG	1799
O	4359	CTAAGCAGAGGCTGCGGTATCATCTGACAAAGTATTCAGCTTGTCTGCATCCAGAC	4418
D	1800	CTAAGCAGTCTGCGGTATCATCTGACAAAGTATTCAGCTTGTCTGCATCCCAAT	1859

OY	4419	TACATGGTCA	CGCCAGGAC	CTAATGTTAC	TACTACTGGCTGGGAGAAACCCAAAGTACC	4478
Db	1860	TATGTGGTCC	TACGACCGAC	CGAATGTTTCA	CTGCTGGGAGAAACCCAAAGTACT	1919
OY	4479	TTTGGACCTG	CGCTTCTCA	GAAGAACCCAG	CTCTTTTGTGAGAAATGAAGTGTGCAT	4538
Db	1920	TTTGGACCTG	CGCTTCTCA	GAAGAACCCAG	CTCTGTGATGAGAAATGAAGTGTGCAT	1979
OY	4539	CACATTAAGT	-----	-----	-----ATATTGTGCTGAGATTTGGCC	4571
Db	1980	CGCATGAGTTT	CTGTAATGGA	AGAAATCCAA	TCCAACTCTGTCTGCGCATTTTGGCC	2039
OY	4572	AGAGGCACTG	ACAGTGGCC	AGGGGTGAC	AGTGGAGGCGCTGTGTTGCTTCAGAGAGAC	4631
Db	2040	GGAGGCACTG	ACAGTGGCC	AGGGGTGAC	AGTGGAGGCGCTGTGTTGCTTCAGAGAGAC	2099
OY	4632	AAATACATTTT	TAACAAGATC	ACTTCTTGGG	GTCTTGGCTGTCAGCCCAATTAAGCT	4691
Db	2100	AAATACATTTT	TAACAAGATC	ACTTCTTGGG	GTCTTGGCTGTCAGCCCAATTAAGCT	2159
OY	4692	GGTGTCTATG	CGTCGTGTTT	CAAGTTTGT	TACTTGAATGAGGAATGATGAGAAATAT	4751
Db	2160	GGTGTCTATG	CGTCGTGTTT	CAAGTTTGT	TACTTGAATGAGGAATGATGAGAAATAT	2219
OY	4752	TAATTGGACG	AGAGACAGAG	CAGAGTGAACA	4778	
Db	2220	TAATTGGACG	AGAGACAGAG	CAGAGTGAACA	2246	
RESULT 11						
AX448883	AX448883	2497 bp	DNA	linear	PAT 03-JUN-2002	
LOCUS	Sequence 1 from Patent WO0220813.					
DEFINITION	AX448883					
ACCESSION	AX448883.1	GI:21697781				
VERSION						
KEYWORDS	human.					
SOURCE	Homo sapiens					
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1					
AUTHORS	Cao, Y.					
TITLE	Materials and methods relating to endothelial cell growth inhibitors					
JOURNAL	Patent: WO 0220813-A 1 14-MAR-2002;					
FEATURES	Karolinska Innovations AB (SE)					
SOURCE	Location/Qualifiers					
	1..2497					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	50..2482					
	/note="unnamed protein product"					
	/codon_start=1					
	/protein_id="CAD37757.1"					
	/db_xref="GI:21697782"					
	/translation="MEHKEVALLLLFLKSGGEPUDYVNTQASLFSYTKQLGAG					
	STEECAACEDEDEFTCRARQYHSKBDQCVIAENRRKSIILIMADVLEKKYVLSE					
	CTKNGKNGRCMTSMKRTNGITLCQKWSSTPHSRRESPTPHSEGLEBNYCRNPNDPQ					
	GWICYTTPDERKRYDYCDILICEECMHGGEYNDIKIKTSGSLCECAMDSQSPHAG					
	YLPSEFNKLKKNYCRNPRLRPMCFCTDDNRMKELDILPRCTPPSSSGPLYCL					
	KQTGNGRYGVAVTVSGHCOHMSAOTPHTRNPRNPENPCNDLENCNRPDKRAMP					
	CHTNSYRPMFYKCTIPSCDSSPYSTQIALPTAPRLPTVNDQCHNGGOSYRGTSSST					
	TTGKACQSWSSMTPHRROKTPENPAPGLTMMYCNRPADKQPMCFPTTDSYVREKCN					
	IAKRGCEASVAPRPVLLPLDVEPTSEEDCKFMGNGKGRKRAATVYGTGPCDDMAO					
	EPHRHSITPETNPRAGLEKNYCRNPDSGVGQPMWYTNPKRLDYDQVQCAAPSD					
	CKRPQVEFKKCPGRVYGVCAVNAHPMPOVSLRTFEFGHFCGGLTISEWVLLPAACL					
	EKSPSSSYKVIILCAHOEVLNLEHVOELIVSRLFLPEPRKIDILKLSASAVLTIVYI					
	PACILSPMYVAADRECFITGMEGQGTGAGAILLEADQVLEKNVCNRPYLENGRVO					
	SEELCACTAGCTPSCDSCGDSGLVPEFKDKIILQGYISWGLCAKRNKCGYVYRISR					
	FTWYIEGYNRN"					

BASE COUNT	692 a	624 c	636 g	545 t
ORIGIN				

Query Match 22.1%; Score 1106.4; DB 6; Length 2497;
 Best Local Similarity 81.4%; Pred. No. 0;
 Matches 1334; Conservative 0; Mismatches 241; Indels 63; Gaps 2;

```

QY 3159 TATCGAGGATATCTCCACCACTGTCAAGAGAGACCTGTCATCTGTCATCTATG 3218
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 896 TATCGGGGGAATGTGGCTCTTACCTGTCCGGGACACCTGTGACACTGGATGACAG 955
QY 3219 ATACCACTGGCATCAGAGAGACCCAGAAAACTAACCAATGTCGCTTGACCGAGAC 3278
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 956 ACCCTTCACACATTAACAGGACACCAAAAACTCCCTGCAAAAAATTTGGATGAAC 1015
QY 3279 TACTGAGGATTCAGATTCCTGGAAACACCTGTGTTACACACCGATCGTGTCG 3338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1016 TACTGCGGCAATCTGACGAAAAAGGGCCCATGTGTCCTACACCAACAGCCAAAGT 1075
QY 3339 AGTGGAGTACTGCAATCTGACACATGCTCAGAAACGAATGATGTCCTAGAGACT 3398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1076 CGGTGGAGTACTGTAAAGTATCCGTCCTG----- 1104
QY 3399 CCCAGCTGTTCACAGTTCACAGAGAGGCTCATTCGAAAGCAGCACCACTGAGCAA 3458
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1105 -TGACTCCCTCCCACTATCCAGAGACAAATGGCTCCCAAGACACCACTGAGCTA 1159
QY 3459 ACCCTGTGTGTCGGCATGCTACCATGTATGSCCAGAGATTTCCAGAGCATTTCTCC 3518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1160 ACCCTGTGTGTCAGAGCATGCTACCATGTATGAGACGACCTCCAGAGCATCTCTCC 1219
QY 3519 ACCACTGTACAGAGAGGACATGTCATCTGTCATCATGACACACACCGGACATAG 3578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1220 ACCACCAACAGAGAAAGATGTCATGTCATCATGACACCAACCGGACACAG 1279
QY 3579 AGGACCCAGAGAAATACCAAAATGATGGCTGACAAATGAACTCTGACAGAAATCCAG 3638
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1280 AAGACCCAGAGAAATACCAAAATGATGGCTGACAAATGAACTCTGAGAAATCCAGAT 1339
QY 3639 GCCGATACAGAGCCCTTGTGTTTACATGAGACCCAGACATCAGATGGAGTACTGCAAC 3698
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1340 GCCGATACAGAGCCCTTGTGTTTACATGAGACCCAGACATCAGATGGAGTACTGCAAC 1399
QY 3699 CTGACGGATGTCAGACAGAGAGGACGTGTCCTCCGACGATGTCATCAGATG 3758
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1400 CTGAAAAAATGCTCAGAGAAACAGAGCGAGTGTAGACCTCCGCCCTGTTCTCTT 1459
QY 3759 CCAAGCCTAGGCTCTCTCTGTAACAAGACTGATGTTGGAAATGGAAAAAGATACCG 3818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1460 CCAAGATGTAGAGACTCTCTCCAGAGAACTGTATGTTGGAAATGGAAAAAGATACCG 1519
QY 3819 GGCAGAGAGGACCACTGTTACTGGAGCGCCATGCCAGAAATGGGCTGCCAGAGCC 3878
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1520 GGCAGAGAGGCGACACTGTACTGGAGCGCCATGCCAGAACTGGGCTGCCAGAGCC 1579
QY 3879 CATAGACAGCAGCTTCATTCAGGAGCAATTAATGGCAGTCTGGAAAAAATTTAC 3938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1580 CATAGACAGCAGCTTCATTCAGGAGCAATTAATTCACGCGGGCTGGAAAAAATTTAC 1639
QY 3939 TGGCGTAACCTGATGATGATCAATCAATGTCCTGCTGCTACACAAATGAAAGAAA 3998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1640 TGGCGTAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1699
QY 3999 CTTTGTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4058
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1700 CTTTGTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1759
QY 4059 CAAATGAGGCGAAGAAATGCTCGAGAGCATGTTAGGGGGGTGTGTGCGCCACCAAT 4118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1760 CAAATGAGGCGAAGAAATGCTCGAGAGCATGTTAGGGGGGTGTGTGCGCCACCAAT 1819
QY 4119 TCTGTGGCGCTGGCAAGTCACTTTAGACAAAGGTTTGAATGCACTTCTGTGAGGAC 4178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1820 TCTGTGGCGCTGGCAAGTCACTTTAGACAAAGGTTTGAATGCACTTCTGTGAGGAC 1879
  
```

```

QY 4179 TTAATATCCCAAGATGGGTGCTGACTGTGCTACCTGCTTGAAGTCTCAAGGCTT 4238
Db 1880 TTGATATCCCAAGATGGGTGCTGACTGTGCTGCTGCTTGGAGAAAGTCCCAAGGCTT 1939
QY 4239 TCATCTACAAAGTCAATCTCTGGTGACACCAAGAAATGAACCTCGAATCTCATGTTCA 4298
Db 1940 TCATCTACAAAGTCAATCTCTGGTGACACCAAGAAATGAATCTCGAATCTCATGTTCA 1999
QY 4299 GAAATGAATGCTGTAGGCTGTCTTGGAGCCCAACAGAGAAATGATGCTGTCTTAAG 4358
Db 2000 GAAATGAATGCTGTAGGCTGTCTTGGAGCCCAACAGAGAAATGATGCTGTCTTAAG 2059
QY 4359 CTAAAGAGGCTCCCTGCTCATCTGACAAAGTAAATGCAAGCTTCTGCTCCACAGAC 4418
Db 2060 CTAAAGAGGCTCCCTGCTCATCTGACAAAGTAAATGCAAGCTTCTGCTCCACAGAC 2119
QY 4419 TACATGCTACCCGACAGACTGAATGTTACATCACTGCTGGGAGAAACCAAGTACC 4478
Db 2120 TATGTGCTGCTGACCGGACCGGATGTTTCATCTGCTGGGAGAAACCAAGTACC 2179
QY 4479 TTTGGGACTGGCTTCCTCAAGAGAGCCAGCTCTTGTATTAAGAAATGAAGTGCAT 4538
Db 2180 TTTGGGACTGGCTTCCTCAAGAGAGCCAGCTCTTGTATTAAGAAATGAAGTGCAT 2239
QY 4539 CACTATTAAGT-----AATTTGCTGGAGCATTTGGCC 4571
Db 2240 CGATATGAGTTTCTGAATGAGAAAGTCCAAATCCACCACTGTGTGCTGGAGCATTTGGCC 2299
QY 4572 AGAGGACATGACAGTTCGCAAGGCTGACAGTGAAGGCTCTGTGTTGCTGAGAGAGAC 4631
Db 2300 GGAGGACATGACAGTTCGCAAGGCTGACAGTGAAGGCTCTGTGTTGCTGAGAGAGAC 2359
QY 4632 AAATCACTTTTACAGAGAGTCACTTCTGGGCTGTGCTGTCGACGCGCCCAATTAAGCT 4691
Db 2360 AAATCACTTTTACAGAGAGTCACTTCTGGGCTGTGCTGTCGACGCGCCCAATTAAGCT 2419
QY 4692 GGTGTCTATGCTGTGCTTTCAGAGGTTGTTACTTGATGATTAAGGAAATGATGAATTAAT 4751
Db 2420 GGTGTCTATGCTGTGCTTTCAGAGGTTGTTACTTGATGATTAAGGAAATGATGAATTAAT 2479
QY 4752 TAATTGAGCGGAGACAG 4769
Db 2480 TAATTGAGCGGAGACAG 2497
  
```

RESULT 12
 AR037325 2497 bp DNA Linear PAT 29-SEP-1999
 LOCUS AR037325
 DEFINITION Sequence 12 from patent US 5801146.
 ACCESSION AR037325
 VERSION AR037325.1 GI:5955181
 KEYWORDS
 SOURCE
 ORGANISM
 UNKNOWN.
 UNCLASSIFIED.
 REFERENCE
 1 (bases 1 to 2497)
 DAVIDSON, D.J.
 TITLE
 JOINTL. Patent: US 5801146-A 12 01-SEP-1998;
 FEATURES
 location/Qualifiers
 1..2497
 /organism="unknown"
 BASE COUNT 691 a 624 c 637 g 545 t
 ORIGIN

Query Match 22.1%; Score 1104.8; DB 6; Length 2497;
 Best Local Similarity 81.4%; Pred. No. 0;
 Matches 1333; Conservative 0; Mismatches 242; Indels 63; Gaps 2;

```

QY 3159 TATGAGGCAATCTCCACCACTGTCAAGAGAGACCTGTCATCTGTCATCTATG 3218
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 896 TATCGGGGGAATGTGGCTCTTACCTGTCCGGGACACCTGTGACACTGGATGACAG 955
  
```

OY	3219	ATACACTGGCATTACAGAGACCACCAAACTACCCAAATGCTGGCTGCACCGAGAC	3278
Db	956	ACCCCTCACACATTAACAGGACACCGAAAATCTCCCTGCCAAAAATTGGATGAAGAAC	10155
OY	3279	TACTGCAGAATTCAGATTCTGGGAAACAACCCGTGATACACACCGATCCGTGTG	3338
Db	1016	TACTGCCGAATTCAGAGCGAAAAAGGCCCATGGTGCATPACCAACGACCCAAAGTG	10757
OY	3339	AGGTGGAGTACTGCATAATCTACACATGCTGTCAAAAACAATCAGGTGTCTTAGAGACT	3398
Db	1076	CGGTGGAGTACTGTAAATACGTCTCTG-----	1104
OY	3399	CCCACTGTGTTCAGTTCCAGATATGGAGGCTCATCTGTGAAGAGACCAATGACGAA	3458
Db	1105	-----TGACTCTCCCCAGTATCCAGCGAATCAATTGGCTCCCACAGCACTGTAGCTA	1159
OY	3459	ACCCCTGTGTCCGCACTGCTACCATGTGAATGCGCAGAGTATCGAGGCACATTTCC	3518
Db	1160	ACCCCTGTGTCCAGACTGCTACCATGTGAATGACAGACTCCAGGACACATCTCC	1219
OY	3519	ACCACTGTACAGGAAGACATGTCAATCTTGATCATCATGACACACACCGCATCAG	3578
Db	1220	ACCACCAACAGAAAGAAATGTCAATGTGGTCTATCATGACACACACCGCACAG	1279
OY	3579	AGGACCCCAAAATACCCCAAATGATGGCTGCACAAATGAATCTCTCAGAGATCCAGAT	3638
Db	1280	AAGACCCCAAAATACCCCAAATGCTGGCTGTACAAATGAATCTCTCAGAGATCCAGAT	1339
OY	3639	GCCGATACAGGCCCTTGTGTATTACATGAGCCACAGCATCAGTGGAGTACTGCAAC	3698
Db	1340	GCCGATTAAGGCCCTGTGTATTACACAGACCCACAGCATCAGTGGAGTACTGCAAC	1399
OY	3699	CTACACCGATGCTGCAGACAGAAAGGACTGTGTGCGCTCCGCTGCATCTCCAGGT	3758
Db	1400	CTGAAAAATGTCTAGGAACAGAAAGGAGTGTATGACACTCCGCTGTGTGCTGCTT	1459
OY	3759	CCAAAGCTAGGCCCTCTTCTGAACAAGACTGTATGTTGGGAATGGGAAAGATACCG	3818
Db	1460	CCAGTATGAGACTCTCTCCGAAAGAGACTGTATGTTGGGAATGGGAAAGATACCGA	1519
OY	3819	GCGAAGAGCAACACTGTACTGGGAGCCAGTGCACAGAAATGGGCTGCCAGAGGCC	3878
Db	1520	GCGAAGAGGCGACCACTGTACTGGGAGCCCATGCACAGACTGGGCTGCCAGAGGCC	1579
OY	3879	CATAGACACAGACCTTATCTCCAGGCAACAATAATGGGAGTCTGSAAAAAATTCAC	3938
Db	1580	CATAGACACAGACTTATCTACCTCGAGACAAATCCACGGGGGCTCTGSAAAAAATTCAC	1639
OY	3939	TGCGGTACCCGTATGTGTGACATCATATGTTCCCTGTGCTACACATGAATCCAGAAA	3998
Db	1640	TGCGGTACCCGTATGTGTGAGTATGAGTGGTCCCTGTGCTACACAAATTCAGAAAAA	1699
OY	3999	CTTTTTGACTGTGTGATATCCCTCTCTGTGCATCCCTTATATTGATTTGGGAAGCTT	4058
Db	1700	CTTTTAGACTATGTGATGTCCCTCATGTGTGGGCCCTCTATTGATTTGGGAAGCTT	1759
OY	4059	CAAGTGACCCGAATAATGTCTCGGAAGATATGAAGGGGGGTGTGGGCCACCCACAT	4118
Db	1760	CAAGTGACCCGAATAATGTCTCGGAAGAGTTGTAGGGGGGTGTGTGGGCCACCCACAT	1819
OY	4119	TGCTGGCCTGGCAAGTCACTCTCAGAACBAAGTTTGGAAAGCACTTGTGGAGSCACC	4178
Db	1820	TGCTGGCCTGGCAAGTCAAGCTTGAAGAACBAAGTTTGGAAAGCACTTGTGGAGSCACC	1879
OY	4179	TTAATVTVCCCAAGTGGGTCTGACAGCTGCTGTCACTCTTGAAGAAGTCTTCAAGGCT	4238
Db	1880	TTGATATVCCCAAGTGGGTGTGACAGCTGCTGTGAAGAGTCTTCAAGAGGCTT	1939
OY	4239	TGATCTTACAAGGTATCTTGGGTGGCAACAAAGTGAATTCGATCTCATTTCTTAC	4298
Db	1940	TGATCTTACAAGGTATCTTGGGTGGCAACAAAGTGAATTCGACCGCATTTCTTAC	1999
OY	4299	GAATVGAAGTCTTAGGCTTTCTTGGAGCCACACAAAGATATTCGCTTGTCAAG	4358

Db	2000	GAATAGAAAGTGTCTAGAGCTTTTGTGGACCCACCAAAAGATATGTGCTTGTGAAG	2059
QY	4359	CTAAGCAGGCGTCCGTCATCTGACAAAGTATGCCAGTGTGTCCATCCGAGC	4418
Db	2060	CTAAGCAGTCTGCGCGTATCTACGACAAAGTAAATCCAGTGTGTGCGCATCCCAAT	2119
QY	4419	TACATGTCACCGCCAGCAGATGTATACATCAGTGTGCTGGGAGAAACCAAGTAC	4478
Db	2120	TATGTGTGCTGACCGGACCGAATGTTGTGTACATGCGCTGGGAGAAACCAAGTACT	2179
QY	4479	TTTGGACGTGCGCTTCTAAGAAAGCCAGCTCTTGTATTATGAAATGAATGTGCAT	4538
Db	2180	TTTGGACGTGCGCTTCTAAGAAAGCCAGCTCTTGTATTGAAATGAATGTGCAT	2239
QY	4539	CACATATAGT-----ATATTGTGCTGAGCATTTGGCC	4571
Db	2240	CGGTATGAGTTCGTAAATGAAAGATCCAAATCCAGCAACTCTGTGTGGGCTTTGGCC	2299
QY	4572	AGAGGACAGACAGTTGCGCAGGAGTACAGGAGGAGCGCTGTGTGCTTCTGAGAAAGAC	4631
Db	2300	GAGGACCTGACAGTTGCGCAGGAGTACAGGAGGAGGAGCTGTGTGCTTCTGAGAAAGAC	2359
QY	4632	AAATACATTTTACAAAGGATCACTTCTTGGGCTTGTGCTGTGCACGCCCAATAAGCT	4691
Db	2360	AAATACATTTTACAAAGGATCACTTCTTGGGCTTGTGCTGTGCACGCCCAATAAGCT	2419
QY	4692	GGGTCTATGTCGCGTTTCAAGTTGTTTACTTGGATTGGAGCAATGATGAGAAATAT	4751
Db	2420	GGGTCTATGTCGCGTTTCAAGTTGTTTACTTGGATTGGAGAGATGATGAGAAATAT	2479
QY	4752	TAAATGACGGAGACAG	4769
Db	2480	TAAATGACGGAGACAG	2497

LOCUS	AR082437	2497 bp	DNA	linear	PAT 31-AUG-2000
DEFINITION	Sequence 12 from patent US 5972896.				
ACCESSION	AR082437				
VERSION	AR082437.1	GI:10009163			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 2497)				
TITLE	Davidson,D.J.				
JOURNAL	Antiangiogenic peptides and methods for inhibiting angiogenesis				
FEATURES	Patent: US 5972896-A 12 26-Oct-1999;				
Source	Location/Qualifiers				
	1..2497				
	/organism="Unknown"				
BASE COUNT	691 a 624 c 637 g 545 t				
ORIGIN					
Query Match	22.1%;	Score 1104.8;	DB 6;	Length 2497;	
Best Local Similarity	81.4%;	Pred. No. 0;			
Matches 1333;	Conservative 0;	Mismatches 242;	Indels 63;	Gaps 2;	
QY 3159	TATGAGGCAATATCCGACACACTGACACAGGAGGACCTGTCAATCTGTGCATATG	3218			
DB 896	TATCGCGGAGATGTGCGTGTACCGTGTCCGGCACACTCTCAGCACCTGGAATGCACG	955			
QY 3219	ATACCACTGGCATCAGAGAGACCCAGAAAACCTAACCCAAATGCTGGCTGACCGAGAAC	3278			
DB 956	ACCCTCACACATTAACAGGACACACAGAAAATCTCCCTGCAAAAATTTTGGATGAAAAC	1015			
QY 3279	TATGCGAGAAATCCAAATTCCTGGGAAACACACCCGGTGTACACAAACCGATCCGTGTG	3338			
DB 1016	TACTGCGCAATCCCTGACGGAAGGAGGCCCATGGTGCCTTACACCAACGACCAATG	1075			
QY 3339	AGGTGGAGTACTGTCAATCTGACACAAATGCTCAGAAACAGAAATCAGGTGTCTAGAGACT	3398			

Db 1076 CGGTGGGAGTACTGTAAAGTACCGTCG----- 1104
QY 3399 CCCAGTGTTCATCTTCCAGCATGAGGCTCATTTGAGCAGCACCACACTGAGCAA 3458
Db 1105 -----TGACGTCCTCCCACTATTCACGGAACAATGGCTCCACAGACACACCTGAGCTA 1159
QY 3459 ACCCTGTGTGTCGGCAGTCTACATGTAATGGCCAGAGTATTTCAGGACCATTTCTCC 3518
Db 1160 ACCCTGTGTGTCGGCAGTCTACATGTAATGGCCAGAGTATTTCAGGACCATTTCTCC 1219
QY 3519 ACCACTGTACAGAGAGACATGTAATCTTGTATCATGACACACACCCGACATG 3578
Db 1220 ACCACACACAGAGAGAGTGTCACTGTGTATGATACACACACCCGACATG 1279
QY 3579 AGAGCCCGAGAGAGAGTGTCACTGTGTATGATACACACACCCGACATG 1339
Db 1280 AAGACCCGAGAGAGAGTGTCACTGTGTATGATACACACACCCGACATG 1399
QY 3639 GCCGATACAGGCTTGTGTGTTTACATGAGACCCGACATCAGGTGGAGTACTGAC 3698
Db 1340 GCCGATACAGGCTTGTGTGTTTACATGAGACCCGACATCAGGTGGAGTACTGAC 1399
QY 3699 CTGACGATGATGTCAGACAGAGAGGAGTGTGTGCTCCGACATGATCAGATG 3758
Db 1400 CTGAAAAATGCTCAGAGACAGAGAGTGTGTGACACCTCCGCTGTGTGTCTGCTT 1459
QY 3759 CCAAGCTTACGCTCTCTCTCTGACACAGATGATGTTGGAAATGGAAAGATACCG 3818
Db 1460 CCAATATGATAGACTCTCTCTCTGACACAGATGATGTTGGAAATGGAAAGATACCG 1519
QY 3819 GGCAGAGAGAGACACATGTTACTGTGAGAGCCATGCCAGATGAGCTGCCAGAGACC 3878
Db 1520 GGCAGAGAGAGACACATGTTACTGTGAGAGCCATGCCAGATGAGCTGCCAGAGACC 1579
QY 3879 CATGAGACAGAGACATGTTACTGTGAGAGCCATGCCAGATGAGCTGCCAGAGACC 3938
Db 1580 CATGAGACAGAGATTTTCTCTGACAGAGAGATGATGTTGGAAATGGAAAGATAC 1639
QY 3939 TGCCGTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3998
Db 1640 TGCCGTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1699
QY 3999 CTTTGTACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4058
Db 1700 CTTTGTACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1759
QY 4059 CAACTGAGGCGGAGAAATGCTCTGAGAGATGATGAGGAGGATGATGAGGCGGACAT 4118
Db 1760 CAACTGAGGCGGAGAAATGCTCTGAGAGATGATGAGGAGGATGATGAGGCGGACAT 1819
QY 4119 TCCGTGGCCCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 4178
Db 1820 TCCGTGGCCCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1879
QY 4179 TTAATATCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4238
Db 1880 TTAATATCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1939
QY 4239 TCATCTACAAAGTCACTCTGAGTGCACACCAAGAGATGATGATGATGATGATGATG 4298
Db 1940 TCATCTACAAAGTCACTCTGAGTGCACACCAAGAGATGATGATGATGATGATGATG 1999
QY 4299 GAAATTAAGTGTCTAGGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 4358
Db 2000 GAAATTAAGTGTCTAGGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 2059
QY 4359 CTAAGCAGGCTGCGCTCATCACTGACAAAGTATGATGATGATGATGATGATGATGATG 4418
Db 2060 CTAAGCAGGCTGCGCTCATCACTGACAAAGTATGATGATGATGATGATGATGATGATG 2119
QY 4419 TACATGCTACCGCAGAGAGTGTATGATGATGATGATGATGATGATGATGATGATGATG 4478

Db 2120 TATGTGTCGCTGACCGGACCGAATGTTTCTGATGCTGCGGAGAACCAAGTACT 2179
QY 4479 TTTGGAGCTGCGCTTCTCAAGAGAGCCAGCTCTTGTATTAAGAAATGAATGTCAT 4538
Db 2180 TTTGGAGCTGCGCTTCTCAAGAGAGCCAGCTCTTGTATTAAGAAATGAATGTCAT 2239
QY 4539 CACTATAGT-----ATATTTGCTGAGCAATTTGGCC 4571
Db 2240 CGTGTATGATTTCTGATGAGAGATCCATTCACCCGACTCTGTGCTGGGCAATTTGGCC 2299
QY 4572 AGAGCAGTACAGTGTCCAGGAGTACAGATGAGAGGCTCTGTGTTCTTCGAGAGAG 4631
Db 2300 GAGGCACTACAGTGTCCAGGAGTACAGATGAGAGGCTCTGTGTTCTTCGAGAGAG 2359
QY 4632 AAATACATTTTACAGAGTCACTCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4691
Db 2360 AAATACATTTTACAGAGTCACTCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2419
QY 4692 GGTGCTATGCTGCTGCTTTCAGAGTGTGTTACTTGTGATGAGAGATGAGAAATAT 4751
Db 2420 GGTGCTATGCTGCTGCTTTCAGAGTGTGTTACTTGTGATGAGAGATGAGAAATAT 2479
QY 4752 TAATTGAGCGGAGACAG 4769
Db 2480 TAATTGAGCGGAGACAG 2497

RESULT 14
AR085163 AR085163 2497 bp DNA linear PAT 01-SEP-2000
DEFINITION Sequence 12 from patent US 5981404.
ACCESSION AR085163
VERSION AR085163.1 GI:10011933
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2497)
AUTHORS Davidson,D.J.
TITLE Antiangiogenic peptides and methods for inhibiting angiogenesis
JOURNALS Patent: US 5981484-A 12 09-NOV-1999;
FEATURES
source 1..2497
BASE COUNT 691 a 624 c 637 g 545 t
ORIGIN

Query Match 22.1%; Score 1104.8; DB 6; Length 2497;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 242; Indels 63; Gaps 2;

QY 3159 TATGAGGCAATATCTTCACACACTGTACAGAGAGACCTGTCAATCTTGGTCAATG 3218
Db 896 TATGAGGCAATATGCTGTATACCGTGTCCGGCACACTGTGACACTGAGATGCACAG 955
QY 3219 ATACCACTGCGATCAGAGAGACCCAGAAAATCTACCAAAATGCTGGCTGACCGAGAC 3278
Db 956 ACCCTCTACACATATACAGAGACACCAAAAATCTCCCTCAAAAATTTGGATGAAAAC 1015
QY 3279 TACTGCAAGATCCAGATTTCTGGAAACAACCTGTTTACACAGACGATCCGTGTGTG 3338
Db 1016 TACTGCGCAATCTCTGACGAAAGAGAGGCGCCATGCTGCTACACACCAACAGCAAGTG 1075
QY 3339 AGTGGAGATATGCAATCTGACACAAATGCTCAGAAAAGATGATGATGATGATGATG 3398
Db 1076 CGGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1104
QY 3399 CCCACTGTTTCTCAGTTCACAGCATGAGGCTCATTTGAGAGACACCACTGAGCAA 3458
Db 1105 -----TGACGTCCTCCCACTATTCACGGAACAATGGCTCCACAGACACACCTGAGCTA 1159
QY 3459 ACCCTGTGTGTCGGCAGTCTACATGTAATGGCCAGAGTATTTCAGGACCATTTCTCC 3518

Db	1160	ACCCTGTGTGTCCAGGACTGTCACATGGTGAATGGACAGACTACGAGGCAATCTCTCC	1212
Oy	3519	ACCACGTCTACAGGAAGGACATGTCAATCTTGTGTATCCATGACACACACGGGATCTACG	3578
Db	1220	ACCAACCAACCAAGGAAGAAAGTGTCACTCTTGTGTATCTTATGACACACACGGGACCGAG	1279
Oy	3579	AGGACCCCGAAAACTACCCCAATGATGGCTGTGACAAATGAATCACTACTCAGAAATCCAGAT	3638
Db	1280	AAGATCCCGAGAAATCTACCCAAATGGCTGTGACTGTGACAAATGAATCACTACTCAGAAATCCAGAT	1339
Oy	3639	GCGCATACAGGCGCTTGTGTGTATTTACCATGAGACCCAGCATCAGTGGGAGTACTGTCAAC	3698
Db	1340	GCGCATTAAGGCGCCCTGTGTGTATTTACCATGAGACCCAGCATCAGTGGGAGTACTGTCAAC	1399
Oy	3699	CTGACCCGATGCTCAGACACAGAAAGGAGTGTGTCTCTCTCCGACTGTGATCCAGGTT	3758
Db	1400	CTGAAAAAATGTCTAGGAACAGAAAGGAGTGTGTGAGACCTCCGCTGTGTGTCTGTGCTT	1459
Oy	3759	CCAAAGCTTGTGGGCGCTCTCTTGTGAACAAAGACTGTATGTTTGGGAATGGGAAGGATACCGG	3818
Db	1460	CCAGATGTGTAGACTCTCTCTCGAAGAAAGACTGTATGTTTGGGAATGGGAAGGATACCGA	1519
Oy	3819	GCGAAGAAAGGCAACACGCTTACTGTGGGAGGAGCCCATCCAGAGAAATGGGTGTGCCAGAGGCC	3878
Db	1520	GCGAAGAAAGGCGACACACTGTACTGTGGGAGGAGCCCATCCAGAGACTGTGGGTGCCAGAGGCC	1579
Oy	3879	CATAGACACAGCACCTTCATTCACAGGAGCAAAATTAATGGGAGGCTGTGAAAAAAATTTAC	3938
Db	1580	CATAGACACAGCATTTTCTTACTACACAGAGCAAAATTCACAGGCGGGGTCTGTGAAAAAAATTTAC	1639
Oy	3939	TGCGCTTAACCCGTGATGGGAGCAATCAATAGGCGCTGTCTACACAAATGAATCCACAAATA	3998
Db	1640	TGCGCTTAACCCGTGATGGGAGCAATGAATAGGCGCTGTCTACACAAATGAATCCACAAATA	1699
Oy	3999	CTTTTGTGACTACTGTGAATACCCCTCTCTGTGCATCTCTTCAATTTGATTTGTGGGAAGCT	4058
Db	1700	CTTTTGTGACTACTGTGAATACCCCTCTCTGTGCATCTCTTCAATTTGATTTGTGGGAAGCT	1758
Oy	4059	CAAGTGAGCCCGCAAGAAATGTCTGTGAAGCAATTTGAGGGGGTGTGTGGCCACCCACAT	4118
Db	1760	CAAGTGAGCCCGCAAGAAATGTCTGTGAAGCAATTTGAGGGGGTGTGTGGCCACCCACAT	1819
Oy	4119	TGCTGTGCGCTGTGCAAGTCAAGTCAAGCAAAAGTTTGAAGAGCACTCTGTGGAGGACAC	4178
Db	1820	TGCTGTGCGCTGTGCAAGTCAAGTCTTAGAACAAAGTTTGAAGAGCACTCTGTGGAGGACAC	1879
Oy	4179	TTTAAATTTCCCAAGAGTGGGTGTGCATGCTGTCTCACTCTTGAAGAAGTCTTAAAGGCT	4238
Db	1880	TTTAAATTTCCCAAGAGTGGGTGTGTACAGTGTGTGCCTCTGTGGAGAAATGTCCTCAAGGCTT	1939
Oy	4239	TCATCTTACAAAGTATCTGTGGGTGCACACCAAGAGTGAACTGTCAATCTCATTTCTTCAG	4298
Db	1940	TCATCTTACAAAGTATCTGTGGGTGCACACCAAGAGTGAAATTCGAAACCGCATGTTTCAG	1999
Oy	4299	GAAATGAGTGTCTAGAGCTGTCTTGTGGAGCCCAACAAGAGATATTTGCTTGTGCTTAAAG	4358
Db	2000	GAAATGAGTGTCTAGAGCTGTCTTGTGGAGCCCAACAAGAGATATTTGCTTGTGCTTAAAG	2059
Oy	4359	CTTAAGCAGGCGTGTGCTATCATCTGACAAAGTAATGCCAGTGTGTCTGTGCTATCCAGAC	4418
Db	2060	CTTAAGCAGTGTGCGGTGTATCTGTACCAAAAGTAATCCAGGTTGTGTGTGCTATCCCAAT	2119
Oy	4419	TACATGTATCCGCGCAGAGACTGATGTATACATCACTGTGCTGTGGGAGAAACCCAAAGTAC	4478
Db	2120	TATGTGTGTGCTGTACCGGACCGAATGTTTGTCTCATGTGCTGTGGGAGAAACCCAAAGTACT	2179
Oy	4479	TTTGGGACGTGGCTTGTCAAGGAAGGCGCAGTCTCTTGTATTTGGAATGAAGTGTGCAT	4538
Db	2180	TTTGGAGCTGTGGCTTGTCAAGGAAGGCGCAGTCTCTGTGATTTGGAATGAAGTGTGCAT	2239
Oy	4539	CACATTAAGT-----ATATTTGTGCTGACATTTGGCC	4571
Db	2240	CGCTTAAGTATTTTCAATGAAGAGTCCATCCACGGAATCTGTCTGTGGGCAATTTTGGCC	2299

QY	4572	AGGCACTGACACTTCCCGAGGGGTGACACTGGAGGGCCCTGCGTTTGGCTTCGAGAAAGAC	4631
DB	2300	GGAGCACTGACACTTCCCGAGGGGTGACACTGGAGGGCCCTGCGTTTGGCTTCGAGAAAGAC	2359
QY	4632	AAATACATTATTCACAAGAGTCACTCTTGGGGCTTTGGCTGACAGCCCCCAATTAAGCTT	4691
DB	2360	AAATACATTATTCACAAGAGTCACTCTTGGGGCTTTGGCTGACAGCCCCCAATTAAGCTT	2419
QY	4692	GCTGCTATGCTGCTGTTTCAAGTTTGTTAAGTTGATTCAGAGGAAATGAGCAAAATAT	4751
DB	2420	GCTGCTATGCTGCTGTTTCAAGTTTGTTAAGTTGATTCAGAGGAAATGAGCAAAATAT	2479
QY	4752	TAATTGACAGGGAGACAG	4769
DB	2480	TAATTGACAGGGAGACAG	2497
RESULT	15		
HOMPLASM		2497 bp mRNA	linear
LOCUS			PRI 27-APR-1993
DEFINITION		Human plasminogen mRNA, complete cds.	
ACCESSION		M74220	
VERSION		M74220.1	GI:190025
KEYWORDS		plasminogen; zymogen.	
SOURCE		Homo sapiens female adult liver CDNA to mRNA.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS		Browne,M.J., Chapman,C.G., Dodd,I., Carey,J.E., Lawrence,G.M.P., Mitchell,I.D. and Robinson,J.H.	
TITLE		Expression of recombinant human plasminogen and aglycoplasminogen in HeLa cells	
JOURNAL		Fibrinolysis (1991) In press	
FEATURES		location/Qualifiers	
SOURCE		1..2497	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/haplotype="diploid"	
		/sex="female"	
		/tissue_type="liver"	
		/dev_stage="adult"	
		50..2482	
		/codon_start=1	
		/product="plasminogen"	
		/protein_id="AAA36451.1"	
		/db_xref="GI:190026"	
		translation="MEHKEVLLLLFLKSGCEPLDYVNTGASLFSYTKKQAGAG SIEEACAKCEDEEPTCAFOYHSKEOQAYIMENKRSIIIMRIVYLFKKYLS CGKGNKNGKMGTSKTKNGITCOKWSSTSPHREAPAHPSGSELENGCRNDPDC GMPCTITDEKRYDYCDLLECEECMCSEENYDGIISKMSGLEQAMDSQSPAHG YPSKPNKNNKNCYCRNRPDELPMCFETDPKRMEDIDIRCTTPPSSGPIYQCI KGTENYRGCVNAVYSGTTCOMASNOTPHNHNPPNPCKNIDNXYCRNDRAPRA CTTNNQVYMEKCTIPSCDSSPYSTBDLATAPELPIYVQCYIHGDSYKGSSTT TTSKROQMSNTPPHRQKTEPNTFNAGLHNTCRNPDADKGCFTTDSYKMECT LKRCSSTESVAPPYVLLPDLVEPTSEECMGNGGTRKATATYGTPODMAAL EPRHSITFETNPRAAGLEKNYCRNPDGSDVGWCTYTNPKLYDCDVPQCAASPS EKSPRPSKYKVLGAHOVNELEPHVOEIEVSRLEPRTDIALTKLSPAVATYAHG SFLCSPPNVVADRTCEYVTGMEGTGCTFGAGILKRAOLPIVENKVCNKEEFLNG PTLCAGHLAGGTCCDGDGSGPLVCFEKRIILIGVTSWGLGCAAPNKPQYVAVS FVYVIGKVRNN"	
		50..106	
		107..2479	
		/product="plasminogen"	
BASE COUNT	691 a	624 c	637 g
ORIGIN		545 t	
Query Match	22.1%	Score 1104.8	DB 9
Best Local Similarity	81.4%	Pred. No. 0	Length 2497
Matches 1333	Conservative	0	Mismatches 242
			Indels 63
			Gaps 2

QY 3159 TATGAGGCATATCTCTCCACACTGTGCACAGAGAGACCTGTCAATCTTGTGTCATGTAG 3218
 Db 896 TATCCGCGGAATGTGGCTGTATCCGTCGCGACACATCTTGTGACATGTCCACAG 955
 QY 3219 ATACCACTGCTGCATCAGAGAGACCCGAGAAATACCCAAATGCTGGCTGCAGAGAAC 3278
 Db 956 ACCCTGCACACATTAACAGAGACACGAAATCTTCCCTCGAAAAATTTGGATGAAAA 1015
 QY 3279 TATGTCAGAGATTCAGANTCTGGGAAACACCTGGGTGTACACACCATCCGTGTGG 3338
 Db 1016 TACTCCGCAATCTGCAGGAAAAAGGCCCCATGGTCATCAACCAACAGCCAAAGTG 1075
 QY 3339 AGGTGGAGTACTGCAATCTGTACACAAATGCTCAGAAAACAGAAATCAGTGTCTAGAGCT 3398
 Db 1076 CGGTGGAGTACTGCAATGATACCTCTCTG----- 1104
 QY 3399 CCCACTGTGTTCAGATTCCAGAGTGAAGGCTCATTTCTGAGACAGCACCAATGAGAA 3458
 Db 1105 -----TGACTCTCCCTCCCGATGCCAGAAACATTTGGCTCCACACACCTGAGCTA 1159
 QY 3459 ACCCTGTGTGTCGCGAGTGTACCATGTATGTGCCAGATTATGAGGCAATTTCTCC 3518
 Db 1160 ACCCTGTGTGTCAGAGTGTACCATGTATGTGACAGAGCTACCGAGCAATCTTCC 1219
 QY 3519 ACCACTGTACAGAGAGACATGTCAATCTTGTGTATGTATGACACACACCGGCACTAG 3578
 Db 1220 ACCACACACAGAGAAAGTGTCAATCTTGTGTATGTATGACACACACCGGCACTAG 1279
 QY 3579 AAGACCCCAAGAAATCTACCAATGTATGCTGTACAAATGAACTACTGTACAGAAATCCAGAT 3638
 Db 1280 AAGACCCCAAGAAATCTACCAATGTATGCTGTACAAATGAACTACTGTACAGAAATCCAGAT 1339
 QY 3639 GCCGATACAGGCTTGTGTGTATTTACATGAGACCCAGCAATGAGGAGAGTACTGCAAC 3698
 Db 1340 GCCGATACAGGCTTGTGTGTATTTACATGAGACCCAGCAATGAGGAGAGTACTGCAAC 1399
 QY 3699 CTGACGCGATGTGCAGACAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3758
 Db 1400 CTGAAAAAATGT 1459
 QY 3759 CCAAGCTTATGAGGCTCTCTTGTGAACAAGACTGTATGTTGGATGGAAAGATACCGG 3818
 Db 1460 CCAAGCTTATGAGGCTCTCTTGTGAACAAGACTGTATGTTGGATGGAAAGATACCGG 1519
 QY 3819 GCGAAGAGGACACACCTGTACTGTGAGAGCCATGCGAGAAATGGCTGCCAGAGAGCC 3878
 Db 1520 GCGAAGAGGACACACCTGTACTGTGAGAGCCATGCGAGAAATGGCTGCCAGAGAGCC 1579
 QY 3879 CATAGACAGACGTCATTCAGAGGACAATTAATGGGACAGTCTGAAAAAATTTAC 3938
 Db 1580 CATAGACAGACGTCATTCAGAGGACAATTAATGGGACAGTCTGAAAAAATTTAC 1639
 QY 3939 TGCCTTAACCTGTATGTGATCAATGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3998
 Db 1640 TGCCTTAACCTGTATGTGATCAATGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1699
 QY 3999 CTTTTTGACTAGTATGATCCCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4058
 Db 1700 CTTTTTGACTAGTATGATCCCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1759
 QY 4059 CAAGTGGAGCCGAAATAATGTCTGTGAAAGCAATTGTAGSGGGGTGTGGCCACCAAT 4118
 Db 1760 CAAGTGGAGCCGAAATAATGTCTGTGAAAGCAATTGTAGSGGGGTGTGGCCACCAAT 1819
 QY 4119 TCTTGCCCTGTGCAAGTCTGTCAAGAAAGATTGTGAAAGCACTTGTGAGAGCACC 4178
 Db 1820 TCTTGCCCTGTGCAAGTCTGTCAAGAAAGATTGTGAAAGCACTTGTGAGAGCACC 1879
 QY 4179 TTAATATCCCGAGATGGGT 4238
 Db 1880 TTAATATCCCGAGATGGGT 1939
 QY 4239 TCAATCTAAGATCATCTGTGGTGTGACACCAAGAAAGTGAACCTGCAATCTATGTTCAG 4298

Db 1940 TCAATCTAAGATCATCTGTGGTGTGACACCAAGAAAGTGAATGTCCAAACGATGTTCAG 1999
 QY 4299 GAAATGAAGTGTCTAGGCTGTCTGTGAGCCACACAGACAGATATTTGCTTTAAG 4358
 Db 2000 GAAATGAAGTGTCTAGGCTGTCTGTGAGCCACACAGAAATATTTGCTTTAAG 2059
 QY 4359 CTAAGAGGCTGTGGCTGTATGATGACAAAGTATGTCACCTGTGTGTCGATCCAGAC 4418
 Db 2060 CTAAGAGTGTCTGGCTGTATGATGACAAAGTATTTCCACCTGTGTGTCGATCCAGAT 2119
 QY 4419 TACATGTACCGCAGAGATGAATGTATCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4478
 Db 2120 TATGT 2179
 QY 4479 TTTGGAGT 4538
 Db 2180 TTTGGAGT 2239
 QY 4539 CACTATAGT-----ATAATTGTGTGATGATTTGGCC 4571
 Db 2240 CGCTATGAGT 2299
 QY 4572 AGAGGACTGACAGT 4631
 Db 2300 GAGGACTGACAGT 2359
 QY 4632 AATATATTTTCAAGAGATCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4691
 Db 2360 AATATATTTTCAAGAGATCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2419
 QY 4692 GGTGTATGTCTGT 4751
 Db 2420 GGTGTATGTCTGT 2479
 QY 4752 TAAATGACGGGAGACAG 4769
 Db 2480 TAAATGACGGGAGACAG 2497

Search completed: March 6, 2003, 12:17:38
 Job time : 12418 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 00:20:02 ; Search time 888 Seconds

Title: US-09-923-515-3_COPY_8939_13938

Sequence: 1 ctccgactgttaccgcggt.....ctgcactatttgcattga 5000

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

```
Minimum DB seq length: 0
```

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*

2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*

3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*

4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*

5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*

6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*

7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1967.DAT.*

8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*

9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*

10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*

11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*

12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*

13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*

14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*

15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*

16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*

17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*

18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*

19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*

20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*

21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*

22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.*

23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001b.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1238.4	24.8	6020	11	AAQ06648	Plasminogen gene f
2	1236.8	24.7	2679	14	AAQ04319	Sequence of a plas
3	1235.4	24.7	2732	20	AX35376	SEQ ID 51 of W0991
4	1235.4	24.7	2732	24	ABN81696	Human Plasminogen
5	1235.4	24.7	6010	12	AAO11998	Human plasminogen
6	1236.8	24.5	2679	18	AAI08686	Plasminogen encodi
7	1234.2	24.5	2753	12	AAO12547	Encodes plasminoge
8	1219.2	24.4	2756	12	AAO12554	Encodes plasminoge
9	1216	24.3	2756	12	AAO12548	Encodes Plasminoge

10	1214.4	24.3	2756	12	AAQ12552	Encodes Plasmid
11	1214.4	24.3	2756	12	AAQ12552	Encodes Plasmid
12	1211.6	24.2	2762	12	AAQ12546	Encodes Plasmid
13	1211.4	24.2	2759	12	AAQ12543	Encodes Plasmid
14	1210	24.2	2762	12	AAQ12544	Encodes Plasmid
15	1210	24.2	2762	12	AAQ12545	Encodes Plasmid
16	1209.6	24.2	2756	12	AAQ12553	Encodes Plasmid
17	1200.2	24.0	2759	12	AAQ12557	Encodes Plasmid
18	1198.6	24.0	2759	12	AAQ12549	Encodes Plasmid
19	1198.6	24.0	2759	12	AAQ12558	Encodes Plasmid
20	1197.8	24.0	2771	12	AAQ12555	Encodes Plasmid
21	1197.4	23.9	2765	12	AAQ12556	Encodes Plasmid
22	1197	23.9	2759	12	AAQ12550	Encodes Plasmid
23	1194.6	23.9	2771	12	AAQ12551	Encodes Plasmid
24	1124.4	22.5	2756	23	AAQ12551	DNA encoding novel
25	1124.4	22.5	2756	23	AAQ12551	DNA encoding novel
26	1111.2	22.2	2286	22	AAQ12747	Plasmid pNZypt-S4
27	1110.6	22.2	2286	14	AAQ040258	Sequence of a DNA
28	1110.6	22.2	2286	21	AAQ080829	Plasmid encoding DNA an
29	1104.8	22.1	2497	21	AAQ052829	Human plasmid
30	1104.8	22.1	2497	24	AAQ05296	Human plasmid
31	1103.2	22.1	2497	24	AAQ05296	Human plasmid
32	1089.8	21.8	2433	20	AAQ177711	Human plasmid
33	1089.8	21.8	2433	24	AAQ08439	Human plasmid
34	919.2	18.4	924	22	AAQ02525	Human angiotensin
35	851.8	17.0	2696	20	AAQ177712	Human angiotensin
36	845.4	16.9	1047	24	AAQ080461	Human miniplasmid
37	721	14.4	2178	23	AAQ080461	DNA encoding novel
38	697.6	14.0	1724	14	AAQ040318	Sequence encoding
39	613	12.3	750	24	AAQ089460	Human microplasmid
40	557.8	11.2	650	15	AAQ053506	Sequence encoding
41	491	9.8	1719	20	AAQ35345	Nucleic acid encod
42	490.6	9.8	1719	20	AAQ35345	Human plasmid
43	429.2	8.6	820	23	AAQ06111	DNA encoding novel
44	429.2	8.6	820	23	AAQ06111	DNA encoding novel
45	311.6	6.2	1353	24	AAQ158257	Canine pro-angiot
ALIGNMENTS						
RESULT 1						
ID	AAQ06648	AAQ06648 standard; DNA; 6020 BP.				
XX	AAQ06648;					
XX	AAQ06648;					
DT	26-FEB-1991	(first entry)				
XX	Plasminogen gene from plasmid p119PN127-6.					
XX	Plasminogen gene from plasmid p119PN127-6.					
XX	Human plasminogen; thrombolytic therapy; p119PN127.6; ds.					
OS	Homo sapiens.					
XX	Homo sapiens.					
FT	Key	Location/Qualifiers				
FT	misc_RNA	1..2294				
FT	misc_RNA	/*tag- a				
FT	insertion_seq	/label=pUC119				
FT	insertion_seq	2295..5098				
FT	insertion_seq	/*tag- b				
FT	insertion_seq	/label=linker				
FT	5'UTR	/note="PMGN CDNA 127"				
FT	5'UTR	2299..2328				
FT	5'UTR	/*tag- c				
FT	CDS	2329..4757				
FT	CDS	/*tag- d				
FT	CDS	/product=plasminogen				
FT	sig_peptide	2329..2385				
FT	sig_peptide	/*tag- e				
FT	mat_peptide	2386..4767				
FT	mat_peptide	/*tag- f				
FT	misc_RNA	2645..2885				

```
FT      /*tag- f
FT      /number=1
FT      /note="k1ingle"
FT      2893..3128
FT      misc_RNA
FT      /*tag- g
FT      /number=2
FT      /note="k1ingle"
FT      3162..3398
FT      misc_RNA
FT      /*tag- h
FT      /number=3
FT      /note="k1ingle"
FT      3468..3704
FT      misc_RNA
FT      /*tag- i
FT      /number=4
FT      /note="k1ingle"
FT      3774..4022
FT      misc_RNA
FT      /*tag- j
FT      /number=5
FT      /note="k1ingle"
FT      4022..4024
FT      misc_RNA
FT      /*tag- k
FT      /note="start" PMGN protein protease"
FT      4074..4106
FT      misc_RNA
FT      /*tag- l
FT      /note="PMGN protein protease activation ARG"
FT      4769..4997
FT      /*tag- m
FT      4997..6020
FT      polyA_site
FT      /*tag- n
XX      W09013640-A.
XX      PN
XX      PD
XX      15-NOV-1990.
XX      PF
XX      26-APR-1990; 90WO-US02296.
XX      PR
XX      01-MAY-1989; 89US-0345801.
XX      PA
XX      (UYNO-) UNIV NOTRE DAME DU.
XX      PI
XX      Caestellino FJ, Whitefleet-Smith JL, Rosen ED, McLinden JH;
XX      WPI: 1990-361475/48.
XX      DR
XX      P-PSDB: NAR08065.
XX      PT
XX      Expression of human plasminogen in eukaryotic cell vector - used
XX      in thrombolytic therapy
XX      PS
XX      Disclosure: Fig 2; 110pp; English.
XX      CC
XX      Eukaryotic cell expression vectors, pref. insect cell expression
XX      vectors, esp. a baculovirus vector, chosen from the gp. of
XX      Autographa californica and Bombyx mori nuclear polyhedrosis virus
XX      CC
XX      reactors, or SV40-, polyoma-, adeno-, VSV or BPV virus, contg. this
XX      CC
XX      gene, can be introduced into hosts, eg. a Spodoptera frugiperda cell.
SQ      Sequence 6020 BP; 1612 A; 1502 C; 1482 G; 1424 T; 0 other:

Query Match      24.8%; Score 1238.4; DB 11; Length 6020;
Best Local Similarity 81.1%; Pred. No. 0;
Matches 1511; Conservative 0; Mismatches 271; Indels 80; Gaps 3;
```

```
OY      3339 AGTGGAGTACTGCAATCTGCACAAATGCTCAGAAAACAGAAATCAGGTGTCTAGAGACT 3398
DB      |||||||
DB      3366 CGGTGGAGTACTGTAAAGATACCGTCTG----- 3394
OY      3399 CCCACTGTTTCCAGTTCCAGTTCGAGGCTCATTTCTGACACACCCAACTGAGCA 3458
DB      ||
DB      3395 -----TGACTCCTCCCAAGTATCCAGGAAACAAATGGCTCCACACACCCACTGAGCTA 3449
OY      3459 ACCCTGTGTCCGGCAGTGTACATGCTAATGGCCAGAGTATTCAGGACATCTTCC 3518
DB      |||||||
DB      3450 ACCCTGTGTCCGGCAGTGTACATGCTAATGGCCAGAGTATTCAGGACATCTTCC 3509
OY      3519 ACCACTGCACAGAGAGACATGCTAATCTTGGTCTCATCTGACACACACCGGCATCG 3578
DB      |||||||
DB      3510 ACCACACACACAGAGAAAGAAATGCACTTGTGGTCACTTGTACACACACCGGCACAG 3569
OY      3579 AGGACCCAGAAAACCTACCCCAATGATGGCTGACATGAATGACTGACAGAAATCCAGAT 3638
DB      |||||||
DB      3570 AAGACCCAGAAAACCTACCCCAATGATGGCTGACATGAATGACTGACAGAAATCCAGAT 3629
OY      3639 GCCGATACAGGCCCTTGGTGTTTTACATGACCCACGATCTCAGGTGGAGTACTGCAC 3698
DB      |||||||
DB      3630 GCCGATTAAGGCCCTTGGTGTTTTACACAGACGCCGCTCAGGTGGAGTACTGCAC 3689
OY      3699 CTGACGGCATGCTCAGACACAGAGGAGACTGTGGTCTCTCCGACTGTCTCAGGTT 3758
DB      |||||||
DB      3690 CTGAAAAAATGCTCAGAGAACAGAGAGCGATGTTGTAGCACTCGCGCTGTCTCTGTT 3749
OY      3759 CCAGGCTAGGGCCTCTTCTGAACAAGACTGTATGTTTGGGAATGGGAAGATATCCG 3818
DB      |||
DB      3750 CCAGATGTAGAGACTCTCTCGAAGAGACTGTATGTTTGGGAATGGGAAGATATCCG 3809
OY      3819 GGCAAGAGGCAACACTGTTACTGAGGAGCCATGCGAGAAATGGGCTGCCAGAGGCC 3878
DB      |||||||
DB      3810 GGCAAGAGGCGACACACTGTTACTGAGGAGCCATGCGAGAAATGGGCTGCCAGAGGCC 3869
OY      3879 CATGACACACACAGCTTCATTCAGGAGCAAAATTAATGGGAGAGTCTGGAAAAATTC 3938
DB      |||||||
DB      3870 CATGACACACACAGCTTCATTCAGGAGCAAAATTAATGGGAGAGTCTGGAAAAATTC 3929
OY      3939 TGCCGTAACCTGATGAGTCAATGATGTCCTGAGTGTCTGATCAAAATGAATCCAGAAA 3998
DB      |||||||
DB      3930 TGCCGTAACCTGATGAGTCAATGATGTCCTGAGTGTCTGATCAAAATGAATCCAGAAA 3989
OY      3999 CTTTTCGACTACTGTGATATCCCTCTCTGTGATCTCTTCATTTGATGTGGAAAGCCT 4058
DB      |||||||
DB      3990 CTTTTCGACTACTGTGATATCCCTCTCTGTGATCTCTTCATTTGATGTGGAAAGCCT 4049
OY      4059 CAAGTGGAGCCGAGAAATGTCTGGAAGCAATGTAGAGGGGTGTGGGCCACCCACAT 4118
DB      |||||||
DB      4050 CAAGTGGAGCCGAGAAATGTCTGGAAGCAATGTAGAGGGGTGTGGGCCACCCACAT 4109
OY      4119 TCCTGGCCCTGGCAAGTCACTGAGCAAGAGTTTGGAAAGCACTTGTGGAGGCACC 4178
DB      |||||||
DB      4110 TCCTGGCCCTGGCAAGTCACTGAGCAAGAGTTTGGAAAGCACTTGTGGAGGCACC 4169
OY      4179 TTAATATCCCGAGAGGGGTGCTGACGTGCTGACGTGTAAGAAAGCTCTGAAGGCT 4238
DB      |||||||
DB      4170 TTAATATCCCGAGAGGGGTGCTGACGTGCTGACGTGTAAGAAAGCTCTGAAGGCT 4229
OY      4239 TCATCTCTCAAGGTTCATCTGGGTGACACCAAGAAAGTGAATCTGAAATCTGTTCC 4298
DB      |||||||
DB      4230 TCATCTCTCAAGGTTCATCTGGGTGACACCAAGAAAGTGAATCTGAAATCTGTTCC 4289
OY      4299 GAAATGAAAGTGTCTAGGCTTTCTTGGAGCCCAACACAGATATGTGCTTGAAG 4358
DB      |||||||
DB      4290 GAAATGAAAGTGTCTAGGCTTTCTTGGAGCCCAACAGATATGTGCTTGAAG 4349
OY      4359 CTAAACGAGCCCTGCGGTATCATGACAAAGTAATGACCTGTCTGACATCCAGAG 4418
DB      |||||||
DB      4350 CTAAACGAGCCCTGCGGTATCATGACAAAGTAATGACCTGTCTGACATCCAGAG 4409
```

[illegible]

PI	Foster DC,	Mulvihill ER,	Ohara PJ,	Pingel K,	Yoshitake S;
XX					
DR	WPI; 1993-133739/16.				
XX	P-PSDB; AAR34428.				
XX					
PT	Human tissue plasminogen activator single chain form fibrinolytic				
PR	agent - comprises thrombin cleavable zymogen stimulating amidolytic activity for lysing clots in heart attack and stroke				
PT	victims and suppressing fibrin matrix				
XX					
PS	Example; Fig 10A, 10B, 10C; 22pb; English.				
CC	A lambda phage clone comprising a cDNA sequence encoding				
CC	plasminogen was obtained from Dr. Mark Marten at the University of				
CC	Washington. The cDNA was isolated from a human liver library by				
CC	probing with the partial sequence of Malinowski et al. The sequence				
CC	of the complete cDNA and the encoded amino acid sequence are shown				
CC	in AAQ40319 and AAR34428.				
XX					
SQ	Sequence 2679 BP; 750 A; 659 C; 675 G; 595 T; 0 other;				
Query Match	24.7%;	Score 1236.8;	DB 14;	Length 2679;	
Best Local Similarity	81.1%;	Pred. No. 0;			
Matches 1510;	Conservative	0;	Mismatches 272;	Indels 80;	Gaps
QY	3159	TATGAGGACATATCTCCACCACACTGTCAAGAAAGACCTTCATTTGGTCACTTANG	3218		
Db	868	TATCGCGGAAGTGTGGCTGTTAACCGTGCCGACCATCTGCAGCACTGACGTGCACAG	927		
QY	3219	ATACCACTGTGCATCAGAGAGACCCAGAAAATTGCCATAATGCTGCCCTGACCGAAGC	3278		
Db	928	ACCCCTACACACATTAACAGAGAACCCAGAAAATTGCCATAATTTGGATGAAAC	987		
QY	3279	TACTGACGAGATCCAAATTCCTGGGAAACAACCCCTGGTCTTACACACCGATCGCTGTG	3338		
Db	988	TATCGCGCAATCTGTACGCGAAAAAGGCGCCATGTGTCATTAACCAACAGCCAAGTG	1047		
QY	3339	AGGTGGAGATCTGCAATCTGCACACANTGCTAGAACAAGATAGGTGTCTTAGAGACT	3398		
Db	1048	CGGTGGAGATCTGAAGTAACTACCGTCTCT-----	1076		
QY	3399	CCCACCTTTTCTCCACTTCCAAAGCATAGAGGCGTCACTTCTGAGCAGCACCATGAGCAA	3458		
Db	1077	-TACTCTCTCCCGAAGTACGAACTAAATGGTCTCCACAGACACCATCTGAGCTA	1131		
QY	3459	ACCCCTGTGTCGGGCAAGTGTACCATGTGTAATGGCCAGATTTATCAGGCAATCTCC	3518		
Db	1132	ACCCCTGTGTCGAGGACGTGTACCATGTGTAATGGGACAGACTVACGAGGACATCTCC	1191		
QY	3519	ACCACTTCACAGAGAGGACATGTCAATTTGNTATCATGACACACACACCGGATCTAG	3578		
Db	1192	ACGACACACAGAAAGAAAGTGTGATGTGTCATGTATGACACACACCGGACACAG	1251		
QY	3579	AGGACCCCAAAAATTACCAATATGTATGGCTTGACATATGAATCTCTCAGAGATCCAGAT	3638		
Db	1252	AAGACCCCAAAAATTACCAATATGTATGGCTTGACATATGAATCTCTCAGAGATCCAGAT	1311		
QY	3639	GCCCATTCAGGCGCTGTGTATTTACATGAGACCCACACATCAGTGGGAGTACTCAAC	3698		
Db	1312	GCGCATTAAGGCGCTGTGTATTTACATGAGACCCACACATCAGTGGGAGTACTCAAC	1371		
QY	3699	CTGACGCGATGTCAAGACAGAAAGGAGCTGTGGTCTCTCCGACTGTTCATCCAGTTT	3758		
Db	1372	CTGAAAAAATGCTACAGAAACAGAAAGGAGTGTGTAGACACTCCGCTGTGTGCTGCTT	1431		
QY	3759	CCAAGCTTAAGGCGCTCTTCTGACACAAAGCTATATTTGGATTTGGAAGATATCCGG	3818		
Db	1432	CCAATGTATAGACTCTCTCCAGAAAGACTGTATGTTTGGATTTGGAAAGATATCCGA	1491		
QY	3819	GCGAAGAGGCACACACTGTATCTGTGGAGCCATCCAGAGATGGCTGCCCGAGGCC	3878		
Db	1492	GCGAAGAGGCACACACTGTATCTGTGGAGCCATCCAGAGATGGCTGCCCGAGAGGCC	1551		

```
QY 3879 CATAGACACGACGCTTCATTCAGGAGCAATAAATGGGCAAGCTCGGAAAAAATTAC 3938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1552 CATAGACACGACATTTCTACTCCAGAGACAATCCAGGGGGGCTGTGGAAAAAATTAC 1611
QY 3939 TCCCTAACCCTGATGGTGACATCATAGTGTCCTGGGCTACACATGAAATCCAGAAAA 3998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1612 TCCCTAACCCTGATGGTGATGTAGGTGTCCTGGGCTACACATGAAATCCAGAAAA 1671
QY 3999 CTTTGTGACTAGTGATATCCCTCTGTGTCATTCCTCTCATTTGATGTGGAGGCT 4058
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1672 CTTTGTGACTAGTGATATCCCTCACTGTGTGGGGCCCTTCATTTGATGTGGAGGCT 1731
QY 4059 CAAGTGGAGCCGAGAAATGTCTGGAGATGTAGGGGGGTGTGGCCCAACCAT 4118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1732 CAAGTGGAGCCGAGAAATGTCTGGAGATGTAGGGGGGTGTGGCCCAACCAT 1791
QY 4119 TCCCTGGCCCTGGCAAGTCAGTCTAGAACAGGTTTGGAAAGCACTTGTGGAGCAC 4178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1792 TCCCTGGCCCTGGCAAGTCAGTCTAGAACAGGTTTGGAAAGCACTTGTGGAGCAC 1851
QY 4179 TTAATATCCCAAGAGTGGGTGCTGACTGCTGCTGGAAGAAGTCTCAAGGCT 4238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1852 TTAATATCCCAAGAGTGGGTGCTGACTGCTGCTGGAAGAAGTCTCAAGGCT 1911
QY 4239 TCATCCCTACAAGTCACTCCGTGGTGCACACCAGAAAGTGAACCTCAATCTGATG 4298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1912 TCATCCCTACAAGTCACTCCGTGGTGCACACCAGAAAGTGAATCTGAACCCGATGT 1971
QY 4299 GAAATAGAAAGTCTAGGCTTCTTGTGAGCCCAACAGACAGATTTGCTGTGAAG 4358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1972 GAAATAGAAAGTCTAGGCTTCTTGTGAGCCCAACAGAAATTTGCTGTGAAG 2031
QY 4359 CTAAACGAGCCCTGCCCTCATCTACAGAACATATGACCACTTGTGCTATCCCAAG 4418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2032 CTAAACGAGCTCGCGCTCATCTACAGAACATATGACCACTTGTGCTATCCCAAG 2091
QY 4419 TACATGTGACCCGACAGACTGATGTATCATCTGCTGGGGAGAAACCAAGGTAC 4478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2092 TATGTGTGCTGCTGACCCGAGCAATGTTCATCATGCTGGGGAGAAACCAAGGTAC 2151
QY 4479 TTTGGAGTGGCTTCTCAAGAGAACCCACCTCTTGTATTTAGAAATGAAGTGCAT 4538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2152 TTTGGAGTGGCTTCTCAAGAGAACCCACCTCTCTGTGATTTAGAAATGAAGTGCAT 2211
QY 4539 CACTATTAAGT-----ATATTTGTCTGACATTTGGCC 4571
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2212 CGCTATGAGTTTGAATGGAAGATCCAAATCCACCGAACTGTGCTGGGCAATTTGGCC 2271
QY 4572 AGAGCACTGACAGTGGCCAGGGGTGACAGTGGAGGGGCTGTGTTGCTTCAGAAAGAC 4631
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2272 GAGGCACTGACAGTGGCCAGGGGTGACAGTGGAGGGGCTGTGTTGCTTCAGAAAGAC 2331
QY 4632 AAATATCATTTTACAGAGAGTCACTTCTGGGGCTTGGCTGTGCACGCCCAATTAAGCT 4691
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2332 AAATATCATTTTACAGAGAGTCACTTCTGGGGCTTGGCTGTGCACGCCCAATTAAGCT 2391
QY 4692 GGTGCTATGCTGTTTCAAGGTTTGTACTTGGATTTGAGGGAAATGTAAGAATAT 4751
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2392 GGTGCTATGCTGTTTCAAGGTTTGTACTTGGATTTGAGGGAAATGTAAGAATAT 2451
QY 4752 TTAATTTGAGGGAGACAGAGTGAAGCATCACTTACTTAGAAGCTGAAGTGGGTAAG 4811
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2452 TTAATTTGAGGGAGACAGAGTGAAGCATCACTTACTTAGAAGCTGAAGTGGGTAAG 2511
QY 4812 ATTTAGCATGTGTAATTAATAGACAGACATCAAGAGACACAGTTCACAGATCAACG 4872
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2512 ATTTAGCATGTGTAATTAATAGACAGACATCAAGAGACACAGTTCACAGATCAACG 2571
QY 4872 CTATGCCAAACCTTGGCAATTTTGTGTATTTTGTGTATTAAGCTTTTAAGTCTGACTGAC 4931
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2572 CTATGCCAAACCTTGGCAATTTTGTGTATTTTGTGTATTTTGTGTATTTTGTGTATTT 2614
QY 4932 AAATTTCTATTTAAGTGTGTATGCTATGCAATTTGTATAAATAAATCACTGCTATAT 4991
```

```
Db 2615 GGAATCTGTAGTAAGGTGACATAGCTATGACATTTGTTAAAAAATCACTGTACTTAAC 2674
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4992 TT 4993
    ||
Db 2675 TT 2676

RESULT 3
AAK35376
ID AAK35376 standard; DNA; 2732 BP.
XX
AC AAK35376;
XX
DT 16-JUL-1999 (first entry)
XX
DE SEQ ID 51 of W09916889.
XX
KW Angiostatin; endostatin; interferon; thrombospondin;
XX
KW Interferon-inducible protein; platelet factor 4; anti-angiogenic;
KW anti-tumor; multifunctional protein; angiogenic-mediated disease;
KW cancer; diabetic retinopathy; macular degeneration; arthritis;
KW tumor cell production; ss.
XX
OS Homo sapiens.
XX
PN W09916889-A1.
XX
PD 08-APR-1999.
XX
PF 30-SEP-1998; 98WO-US20464.
XX
PR 01-OCT-1997; 97US-0060609.
XX
PA (SEAR ) SEARLE & CO G D.
XX
PI Bolanowski MA, Caparon ME, Casperson GF, Gregory SA;
PI Klein BK, McKearn JP;
XX
DR WPI; 1999-255098/21.
XX
PT New multifunctional proteins useful for treating angiogenic-mediated
XX
PT diseases
XX
PS disclosure; Page 85-86; 121pp; English.
XX
CC The specification describes multifunctional proteins which comprise
CC combinations of angiostatin, endostatin, interferon, thrombospondin,
CC interferon-inducible protein and platelet factor 4, and have
CC anti-angiogenic and/or anti-tumor activity. The multifunctional protein
CC may exhibit useful properties such as having similar or greater
CC biological activity when compared to a single factor or by having
CC improved half-life or decreased adverse side effects, or a combination
CC of these properties. The proteins can be used for treating an
CC angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular
CC degeneration, or arthritis. They can also be used for inhibiting the
CC production of tumor cells (characteristic of lung, breast, ovarian,
CC prostate, pancreatic, gastric, colon, renal, bladder cancers; melanoma,
CC hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor
CC growth. The present sequence is used in the course of the invention.
XX
SQ Sequence 2732 BP; 757 A; 667 C; 690 G; 618 T; 0 other;

Query Match 24.7%; Score 1235.4; DB 20; Length 2732;
Best Local Similarity 80.9%; Pred. No 0;
Matches 1511; Conservative 0; Mismatches 276; Indels 80; Gaps 3;

QY 3159 TATCGAGCATATCCCTCACACACATGTCACAGAGACCTGTCAATCTTGGTCACTATG 3218
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 901 TATCGGGGGAATGTGCTGTACCGTGTCCGGCACACCTGTGACGACGTGAGATGACAG 960
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3219 ATACCACTGTGCATCAGAGACCCAGAAAAATCAACCAATGCTGGCTGACCGGAGAC 3278
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 961 ACCCCACACACATTAACAGGACACAGAAAATTCCCTGCAAAAATTTGATGAAAC 1020
QY 3279 TACGTGAGGAATCCAGATTTCTGGGAAACAACCTGGGTGTTACACACCGATCCGTGTG 3338
Db 1021 TACTGCCGAAATCTGAGGAAAAAGGGCCCTGATGCCATTCACCAACCAAGG 1080
QY 3339 AGGTGGAGTACTGCAATCTGACAAATGCTCAGAAAACAGAAATCAGGTGCTTAGAGACT 3398
Db 1081 CGGTGGAGTACTGTAAGATACCGTCTG----- 1109
QY 3399 CCCAGTGTCTCCAGTTCGAAGCATGAGAGCTCATTTGAGAGCACCAACTGAGCAA 3458
Db 1110 -----TGACTCTCCCGCAGTATCCAGGACAGATTGGCTCCACAGACCACTGAGCTA 1164
QY 3459 ACCCTGTGGTCCGCGAGTGTACCATGATGTAATGGCCAGAGTTATCGAGCAGATTCCTC 3518
Db 1165 ACCCTGTGGTCCGAGACTGCTACCATGATGATGACAGAGCTACCGAGCAGATTCCTCC 1224
QY 3519 ACCACTGTACAGAAAGACATGTCAATCTGTGTCATTCATGACACACACCGGATCAG 3578
Db 1225 ACCACACACAGAAAGAGTCAAGTGTGTCATCTATGACACACACCGGACACAG 1284
QY 3579 AGGACCCCGAGAAATACCAATGATGCTGACAAATGAAGTACTGACGAATCCAGAT 3638
Db 1285 AAGACCCCGAGAAATACCAATGATGCTGACAAATGAAGTACTGACGAATCCAGAT 1344
QY 3639 GCGGATACAGGCCCTTGGTGTTTTACCATGATGACCCAGCATCAGGTGGAGTACTGCAAC 3698
Db 1345 GCGGATTAAGGCCCTTGGTGTTTTACCAAGACCCAGGCTCAGAGTGGAGTACTGCAAC 1404
QY 3699 CTGACGCGATGCTCAGACACAGAAAGGAGTGTGCTGCTCCGATGTCATCAGGT 3758
Db 1405 CTGAAAAATGCTCAGGAAACAGAAACGAGTGTGAGACACTCCGCTGTGTCCTGCTT 1464
QY 3759 CCAAGGCTAGGGCTCTCTGGAACAAGATGTAATGTTGGGAAATGGGAAAGATACCG 3818
Db 1465 CCAAGTGTAGAGACTCTCTCCAGAAAGACTGTATGTTGGGAAATGGGAAAGATACCGA 1524
QY 3819 GGCAGAAAGCAACCACTGTACTGTGGAGGCCATGCCAGATGGGCTGCCAGAGCCCC 3878
Db 1525 GGCAGAAAGGCGCACACTGTACTGTGGAGGCCATGCCAGAGCTGGCTGCCAGAGCCCC 1584
QY 3879 CATAGACAGGACGCTTACTTCCAGGACCAATAATAGGAGGATCGGAAAAAATATAC 3938
Db 1585 CATAGACAGGACGCTTACTTCCAGGACCAATAATAGGAGGATCGGAAAAAATATAC 1644
QY 3939 TGCCTTAACCTGTATGTGACATCAATGTCCCTGTGTCTACACAAATGAATCCAGAAAA 3998
Db 1645 TGCCTTAACCTGTATGTGATGTAGTGTCCCTGTGTCTACACAAATCCAGAAAA 1704
QY 3999 CTTTTGATTAATGATATCCCTCTCTGTGATCTCTTATTTGATTTGGGAAAGCT 4058
Db 1705 CTTTTGATTAATGATATCCCTCTCTGTGATCTCTTATTTGATTTGGGAAAGCT 1764
QY 4059 CAAGTGGAGCCGAGAAATGTCCTGGAAGCATTTGAGGGGGGTGTGGCCACCCACAT 4118
Db 1765 CAAGTGGAGCCGAGAAATGTCCTGGAAGGGTGTGGGGGGGTGTGGCCACCCACAT 1824
QY 4119 TCTGGCCCTGCGAAGTCTCAGAACCAAGGTTTGGAAAGCACTTCTGTGGAGGACC 4178
Db 1825 TCTGGCCCTGCGAAGTCTCAGAACCAAGGTTTGGAAAGCACTTCTGTGGAGGACC 1884
QY 4179 TTAATTTCCCGAGAGTGTGCTGACTGCTGCTCAGTCTCTGAGAGAGTCTCTCAAGCCT 4238
Db 1885 TTAATTTCCCGAGAGTGTGCTGACTGCTGCTCAGTCTCTGAGAGAGTCTCTCAAGCCT 1944
QY 4239 TCAATCTAAGAGTCAATCTGTGGTGCACACCAAGAGTGAACCTCGATCTCATGTTCA 4298
Db 1945 TCAATCTAAGAGTCAATCTGTGGTGCACACCAAGAGTGAATCTCGAACCGCATGTTCA 2004
QY 4299 GAAATAGAGTGTAGGCTGTCTTGAGAGCCACACCAAGAGATATTGCTCTTAAG 4358
Db 2005 GAAATAGAGTGTAGGCTGTCTTGAGAGCCACACCAAGAAAGATATTGCTCTTAAG 2064

QY 4359 CTAGCAGGCCCTGCCGATCACTAGCAAGTATGCCAGGCTGTGCTGCATCCCGAGAC 4418
Db 2065 CTAGCAGGCTCCTGCTGATCACTGACAAAGTAAATCCAGCTGTCTGCTATCCCAAT 2124
QY 4419 TACATGTCACCCGAGAGTGAATGTTACATCACTGCTGGGGAGAAACCAAGATACC 4478
Db 2125 TATGTGCTCCGTCAGCCGAGCCGAAATGTTTCACTACCTGGCTGGGGAGAAACCAAGATACC 2184
QY 4479 TTTGGAGCTGCCCTTCTCAAGAAAGCCAGGCTCCTGTTATTTAGAAATGAAGTGTCAAT 4538
Db 2185 TTTGGAGCTGCCCTTCTCAAGAAAGCCAGGCTCCTGTTATTTAGAAATGAAGTGTCAAT 2244
QY 4539 CACTATAAGT-----ATATTTGTGTCAGCAATTTGGCC 4571
Db 2245 CGTATGAGTTTCTGAAATGAGAAAGTCCATACCCAGAACTCTGTGCTGGGCAATTTGGCC 2304
QY 4572 AGAGGCACTGACAGTTGCCAGGGGTGACAGTGAAGGGCCCTGGTGTGCTTGGAGAGAGAC 4631
Db 2305 GGAGGCACTGACAGTTGCCAGGGGTGACAGTGAAGGTCCTGTGTTGCTTGGAGAGAGAC 2364
QY 4632 AAATACATTTTAAAGAGTCACTCTTGGGGCTTGGCTGTGCACGCCCAATTAAGCCT 4691
Db 2365 AAATACATTTTAAAGAGTCACTCTTGGGGCTTGGCTGTGCACGCCCAATTAAGCCT 2424
QY 4692 GGTGCTATGCTGCTGTTTCAAGTGTGTTACTGTGATTTGAGGAGGAAATGAGAAATAT 4751
Db 2425 GGTGCTATGCTGCTGTTTCAAGTGTGTTACTGTGATTTGAGGAGGAAATGAGAAATAT 2484
QY 4752 TAAATGGACGGAGACAGAGTGAAGCATCACTTCTTAAAGCTGAAGAACTGGGTTAAG 4811
Db 2485 TAAATGGACGGAGACAGAGTGAAGCATCACTTCTTAAAGCTGAAGAACTGGGTTAAG 2544
QY 4812 AATTAGCATGCTGGAATTAATAGACAGATCAACAGAAAGACAGTCCAGGTACAG 4871
Db 2545 AATTAGCATGCTGGAATTAATAGCTGATATCAACAGAAAGACAGTCCAGGTACAG 2604
QY 4872 CTATGCCAAACCTTGCCATTTTGGTATTTTGTGTATTAAGCTTTAAGGTCTGACAG 4931
Db 2605 CTATGCCAAACCTTGCCATTTTGGTATTTTGTGTATTT-----TTCTGACTGCT 2647
QY 4932 AATTCTGTAATAGTGTATAGTATGACATTTGTTAAAAATTAACCTGCACCTATT 4991
Db 2648 GGATTCGTGTAAGTGAAGTGAATAGTATGACATTTGTTAAAAATTAACCTGCTACTTAC 2707
QY 4992 TTGATTT 4998
Db 2708 TTGATTT 2714

RESULT 4
ID ABN81696 standard; DNA: 2732 BP.
XX ABN81696;
DT 03-SEP-2002 (first entry)
XX
DE Human plasmidogen encoding DNA.
XX
KW Human; plasmidogen; lvs plasmidogen; glu plasmidogen; A61; p22;
KW vascular endothelial; cell proliferation; anti-angiogenic; cancer;
KW rheumatoid arthritis; Crohn's disease; diabetic retinopathy; anti-cancer;
KW cytostatic; gynaecological; neuroprotective; antipsoriatic;
KW antiarteriosclerotic; dermatological; antineumatic; antiarthritis;
KW antiinflammatory; antidiabetic; ophthalmological; immunosuppressive;
KW cardiac; vulnerary; vasotropic; anti-tumour; gene; ds.
XX
OS Homo sapiens.
XX
XX
XX
XX Key Location/Qualifiers
FH 1..2487
FT CDS
FT /*tag= a

FT			/product= "plasminogen"
FT	sig_peptide	55..111	/*lag- b
FT		112..2484	/*lag- c
FT	mat_peptide	/product=	"glu plasminogen"
FT		343..2484	/*tag= d
FT	mat_peptide	/product=	"lys plasminogen"
FT		343..1515	/*lag= e
FT	mat_peptide	/product=	"A61 isoform"
FT		343..1524	/*tag= f
FT	mat_peptide	/product=	"A61 isoform"
FT		343..651	/*lag= g
FT		/product=	"P22"
XX			
PN	WO200244328-A2.		
XX			
PD	06-JUN-2002.		
XX			
PR	28-NOV-2001; 2001MO-US44515.		
XX			
PA	(NAIS/) WAISMAN D M.		
PA	(KASS/) KASSAM G.		
XX	(KWON/) KWON M.		
PI	Waisman DM, Kassam G, Kwon M;		
XX			
DR	WPI: 2002-527706/56.		
DR	P-PSDB: ABB83795.		
XX			
PT	Novel naturally occurring fragment A16 or P22 of plasminogen, useful		
PT	for treating for anti-angiogenic treatment of a mammal suffering from		
PT	cancer and inhibiting proliferation of vascular endothelial cells -		
XX			
PS	Claim 26; Fig 2; 88pp; English.		
XX			
CC	The invention relates to an isolated naturally occurring fragment A61		
CC	or p22 polypeptide (I) of plasminogen or an isolated polypeptide		
CC	comprising 103, 391 or 394 contiguous amino acids of amino terminal of		
CC	plasminogen (ABB83795). (I) has vascular endothelial cell proliferation		
CC	inhibitor activity. (I) is useful for anti-angiogenic treatment of a		
CC	mammal suffering from cancer, acoustic neuromas, neurofibromas,		
CC	atherosclerosis, rheumatoid arthritis, Crohn's disease, scleroderma,		
CC	adiposity, pyogenic granuloma, rubecosis, and diabetic retinopathy,		
CC	retinopathy of prematurity, neovascular glaucoma, retrolental		
CC	fibroplasia, graft rejection, myocardial angiogenesis, plaque		
CC	neovascularization, haemophilic joints, angiodroma and wound		
CC	granulation and inhibiting proliferation of vascular endothelial cells,		
CC	in combination with an anti-cancer agent e.g. methotrexate, mitozantrone,		
CC	palliative, vinblastine, 5-fluorouracil, cisplatin, leucovorin,		
CC	cyclophosphamide and oncolytic virus.		
XX			
SQ	Sequence 2732 BP; 758 A; 669 C; 688 G; 617 T; 0 other:		
	Query Match	24.7%	Score 1235.4; DB 24; Length 2732;
	Best Local Similarity	80.9%;	Pred. No. 0;
	Matches 1511; Conservative	0;	Mismatches 276; Indels 80; Gaps 3
OY	3159 TATGAGGCATATCTCCACCACTGTCCACAGAAGACCTTCATCGTGGTCATGTATG		3218
DB	901 TATGGCGGAATGTGCTGTACCGTTGCCGGCACACTGTCAAGTGAGTGCACAG		960
OY	3219 ATACACACATCGCATAGAGGACCCAGAAACTRACCCAATTCGTGGCCTGACCAGAAC		3278
DB	961 ACCCTTCACACATTAACGAGACACCGAAAACCTCCCTCCAAAAAATTGATGAATAAC		1020

QY	3279	TACTGCAGGATTCAGGATTTCTGGGAAACAACCTGGTGTTTACACACCGATCCGTGTGG	3336
Db	1021	TACTGCAGGATTCAGGATTTCTGGGAAAGAGGGCCCTGGTGTTCATACACCGATCCAGCGATG	1080
QY	3339	AGGAGGAGATTCGGCATTCCTACACCAATGGCTGAGAAAGAAATCAGAGGTGTCTAGAGACT	3398
Db	1081	CGGAGGAGATTCGTCATACATACCTCTCTG-----	1105
QY	3399	CCCACGTGTGTTCAGGTTCCAGGATCAGGAGGCTCATTTCTGAAGCAGCACCACTGAGCAA	3455
Db	1110	-----TGACTCCCTCCCAAGTATCCACGGAAACAATGGCTCCACAGCACCACTGTAGCTA	1164
QY	3459	ACCCGATGGATCCGGCAGTGTCTACCATGGTAAATGGCCANAGTTATTCAGGACCATTTCTCC	3518
Db	1165	ACCCGATGGATCCAGGATGTCTACCATGGTAAATGGCCANAGTTATTCAGGACCATTTCTCC	1222
QY	3519	ACCACTGTACAGGAGAGACATGTCACTCTTGGTCAATCCATACACACACCGGGCATCAG	3578
Db	1225	ACCACTGTACAGGAGAGAGATGTCACTCTTGGTCAATCCATACACACACCGGGCATCAG	1284
QY	3579	AGGACCCCAAGAAATATACCCAAATGATGGGCTACATAGAACTACTCAGAGAAATCCAGAT	3633
Db	1285	AAGGACCCCAAGAAATATACCCAAATGATGGGCTACATAGAACTACTCAGAGAAATCCAGAT	1344
QY	3639	GCCGATTCAGGCTCTTGGTGTTTACCATGTGACCCAGCATATCAGTGGGAGTCTCTCAAC	3698
Db	1345	GCCGATTCAGGCTCTTGGTGTTTACCATGTGATTCACACAGACCCAGCTACAGTGGGAGTCTCTCAAC	1404
QY	3699	CTGACCCGATGTCTGACACACAGGAGGAGCTGTGGTGGCTCTCCGATCTGCATCAGAGTT	3758
Db	1405	CTGACCAAAATGTCTGAGAAACAGAAAGGAGTGTGTAGCACCTCCGCTGTGTCTGTCTT	1464
QY	3759	CCAAAGCTAGGGCCCTCTCTGAAACAAGACTGTATGTTGGAAATGGGAAAGGATACCG	3818
Db	1465	CCGATGTGTGAAGCTCTCTCGAAGAAAGTGTATGTTGGAAATGGGAAAGGATACCGA	1524
QY	3819	GGCAGAAAGGCAACCACTGTACTGAGGAGCCCATGCGCAGGATGGCTGCCAGAGGCC	3878
Db	1525	GGCAGAAAGGCGCACCACTGTACTGAGGAGCCCATGCGCAGGATGGCTGCCAGAGGCC	1584
QY	3879	CATAGACACAGCAGCTTATTCACAGGGGACAAATTAATGGGAGGCTCGGAAAAATTAAC	3938
Db	1585	CATAGACACAGCAGATTTTACCTCCAGAGACAAATCCACGGGGGCTCGGAAAAATTAAC	1644
QY	3939	TGCCGTAAACCTGATGTGATCATCAATGTCCTGGTGTCTACACAAATGAATCCAGAAAA	3998
Db	1645	TGCCGTAAACCTGATGTGATGTGATGAGTGTGGTGTCTGTCTACACAAATGAATCCAGAAAA	1704
QY	3999	CTTTTGAATCTGTGATATCCCTCTCTGTCGATCCTCTTATTTGATTTGGAAAGCCT	4058
Db	1705	CTTTTGAATCTGTGATATCCCTCTCTGTCGATCCTCTTATTTGATTTGGAAAGCCT	1764
QY	4059	CAGGTGAGGACCGAAGAAATGCTCTGGAACATTTGAAGGGGGGTGTGGCCACCCACAT	4118
Db	1765	CAGGTGAGGACCGAAGAAATGCTCTGGAAGGGGTGTGGGGGGGTGTGGGGCCACCCACAT	1824
QY	4119	TCCGTGGCCCTGTGCAAGTCACTGTCAAGAAACAGTTTGGAAAGCACTCTTGGAGGACAC	4178
Db	1825	TCCGTGGCCCTGTGCAAGTCACTGTGCAAGAAAGTTTGGAAATGCACTCTTGGAGGACAC	1884
QY	4179	TTAAATATCCCAAGGTGGGTGTGATCTGTCTGCTCACTCTTGAAGAGTCTCAAGGCT	4238
Db	1885	TTAAATATCCCAAGGTGGGTGTGATCTGTCTGCTCACTCTTGAAGAGTCTCAAGGCT	1944
QY	4239	TCAATCTTACAGGTTCATCTGGGTGTGACACACAAAGATGAACCTCGAATCTCATTTTTCAG	4298
Db	1945	TCAATCTTACAGGTTCATCTGGGTGTGATCTGTGACACACAAAGATTCGAAACGATGTTCAG	2004
QY	4299	GAAATATGAAGTCTCAGGCTGTCTTGGAGGCCACAGAAAGATATGTGCTTGGCTAAAG	4358
Db	2005	GAAATATGAAGTCTCAGGCTGTCTTGGAGGCCACAGAAAGATATGTGCTTGGCTAAAG	2064
QY	4359	CTTAAGAGGCTGCCGTTCATCACTGACAAAGTAATGTCAGGTGTGTGTCATCTCCCAAG	4418

```
Db 2065 CTAAGCAGTCCGCGCTCATCTGACAAAGTATCCACGCTTGCTGCAATCCCAAT 2124
OY 4419 TACATGGTCACCCGAGACTGATGTACATCTGCTGGGAGAAACCAAGTACC 4478
Db 2125 TATGTGGTCCGTGACCGGACGGAATGTTTCATCAGCTGGGAGAAACCAAGTACT 2184
OY 4479 TTTGGAGCTGGCCCTTCACAGGAAGCCAGCTCCTGATTGAGAAATGAGTGCAT 4538
Db 2185 TTTGGAGCTGGCCCTTCACAGGAAGCCAGCTCCTGATTGAGAAATGAGTGCAT 2244
OY 4539 CACTATAGT-----ATATTGGCTGAGCATTTGGCC 4571
Db 2245 CGCTATAGGTTTGAATGGAAGATCCATCCACCGAACTGCTGGCATTTGGCC 2304
OY 4572 AGAGGCACTGACAGTTGCCAGGGTGACAGTGAAGGCGCTCTGTTGCTTCAGAGAGAC 4631
Db 2305 GGAGGCACTGACAGTTGCCAGGGTGACAGTGAAGGCGCTCTGTTGCTTCAGAGAGAC 2364
OY 4632 AAATACATTTTACAGAGAGTCACTCTTGGGGTCTTGCTGTCACGCCCCAATAAGCT 4691
Db 2365 AAATACATTTTACAGAGAGTCACTCTTGGGGTCTTGCTGTCACGCCCCAATAAGCT 2424
OY 4692 GGTGCTATGCTGCTGTTTCAAGTTTGATCTTGATGAGAGGAATGATGAATAAT 4751
Db 2425 GGTGCTATGCTGCTGTTTCAAGTTTGATCTTGATGAGAGGAATGATGAATAAT 2484
OY 4752 TAAATGACGGGAGACAGAGTGAACATCACTACTTGAAGCTGAAAGCTGGGTAAG 4811
Db 2485 TAAATGACGGGAGACAGAGTGAACATCACTACTTGAAGCTGGGTAAG 2544
OY 4812 ATTTAGCATGCTGGAATTAATAGACGATCAAGCAAGACACTGTTCCAGTACCAG 4871
Db 2545 ATTTAGCATGCTGGAATTAATAGACGATCAAGCAAGCAAGTCCCGACTACAG 2604
OY 4872 CTATGCCCAACCTTGGCATTTTGGTATTTTGTATAGCTTTTAAGGTGACTGAC 4931
Db 2605 CTATGCCCAACCTTGGCATTTTGGTATTTTGTATAGCTTTTGTATAGCT 2647
OY 4932 AAATCTGATTAAAGTGTATAGTATGATCAATTTGTTAAATAAATCACTGCACTTAT 4991
Db 2648 GGATTCGTAGTAAAGTGACATAGCTATGACATTTGTTAAATAAATCACTGCACTTAC 2707
OY 4992 TTGATTT 4998
Db 2708 TTGATTT 2714

RESULT 5
AAQ11998
ID AAQ11998 standard; DNA: 6010 BP.
XX
AC AAQ11998;
XX
DT 05-SEP-1991 (first entry)
XX
DE Human plasminogen with substd. Arg561 from pUC119PM127.6 vector.
XX
KM Plasminogen; proteolytic cleavage; variant; thrombosis; HPg; ds.
XX
OS Homo sapiens.
XX
FH Key CDS location/Qualifiers
FT CDS 2329..4758 /product= a
FT sig_peptide 2329..2385 /product= plasminogen variant
FT mat_peptide 2386..4758 /tag= b
FT 2386..4758 /tag= c
FT misc_recomb 2284
FT 2284
FT /tag= d
FT /note= "end pUC119 vector"
```

```
FT misc_recomb 2285
FT /tag= e
FT /note= "pmgn cDNA linker start"
FT 5'UTR 2294
FT /tag= f
FT /note= "pmgn mRNA 5'UT start"
FT misc_RNA 2635..2874
FT /tag= g
FT /label= ktingle_1
FT 2881..3117
FT /tag= h
FT /label= ktingle_2
FT 3151..3387
FT /tag= i
FT /label= ktingle_3
FT 3457..3693
FT /tag= j
FT /label= ktingle_4
FT 3769..4011
FT /tag= k
FT /label= ktingle_5
FT misc_RNA 4009
FT /tag= l
FT /note= "start pmgn protein protease"
FT 3'UTR 4759
FT /tag= m
FT /note= "start pmgn mRNA 3'UT"
FT 4986..5074
FT /tag= n
FT /label= poly(A)+
FT misc_recomb 5087
FT /tag= o
FT /note= "start pUC119 vector and end cDNA 127 linker"
FT 4066..4068
FT /tag= p
FT /label= mutation site
FT /note= "wild-type: protease activation Arg"
XX
PN W09108297-A.
XX
PD 13-JUN-1991.
XX
PF 31-OCT-1990; 90WO-US06345.
XX
PR 01-DEC-1989; 89US-0444584.
XX
PA (GENE ) GENENTECH INC.
XX
PI Castellino FJ, Higgins D L;
XX
DR WPI, 1991-193201/26.
XX
DR P-PSDB; AARI2406, AARI3219, AARI3220.
XX
PT New human plasminogen variant with replaced ARG-561 - is
PT complexed with fibrinolytic enzyme for use in thrombolytic
PT therapy
XX
PS Disclosure: Fig 1(A-P); 82pp; English.
XX
CC The nucleotide at position 3809 is T for the vector pUC119PM127.6,
CC whereas Fig 1 of the specification reflecting the sequence
CC encoding native human plasminogen, contains a C. The codons at
CC positions 4066-4068 (NNN) which encode the amino acid Arg561 in the
CC wild-type plasminogen, are substituted to encode any amino acid
CC except Lys. Pref. the codon encodes amino acids Ser, Gly or Glu.
CC Position 561 is the critical cleavage site in the conversion of
CC plasminogen to plasmin. The resulting product is proteolytic
CC resistant plasminogen which may be used to treat thrombosis in
CC humans. When complexed with streptokinase, it does not degrade into
CC plasmin and it caused rapid plasminogen activator activity.
XX
SQ Sequence 6010 BP; 1611 A; 1497 C; 1475 G; 1424 T; 3 other;
```


OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	22..2454
FT	/tag= a
FT	/transl_except= (pos 199..201, aa: Tyr)
FT	/transl_except- (pos 592..594, aa: Gly)
FT	/transl_except- (pos 688..690, aa: Cys)
FT	/transl_except- (pos 859..861, aa: Gly)
FT	/transl_except- (pos 892..894, aa: Val)
FT	/transl_except- (pos 1528..1530, aa: Glu)
FT	/transl_except= (pos 1840..1842, aa: Cys)
FT	/transl_except= (pos 1864..1866, aa: Glu)
FT	/transl_except- (pos 1993..1995, aa: Phe)
FT	/transl_except- (pos 2146..2148, aa: Gly)
FT	/note= "encodes protein plasminogen"
FT	unsure
FT	1304...1206
FT	/tag= b
FT	/note= "encodes gly"
XX	
PN	US5648254-A.
XX	
PD	15-JUL-1997.
XX	
PF	14-JUL-1994;
XX	
PR	04-DEC-1989;
XX	
PR	15-JAN-1988;
XX	
PR	28-OCT-1991;
XX	
PR	14-JUL-1994;
XX	
PA	(ZYMO) ZYMOGENETICS INC.
PI	Kumar AA, Mulvihill ER;
XX	
DR	WPI: 1997-372063/34.
XX	
PT	P-PADB: AAWJ1169.
XX	
PT	Production of recombinant plasminogen - by co-expression with
XX	
PT	plasminogen-processing or -stabilising protein
XX	
PS	Example 2; Fig 6A-D; 32pe: English.
XX	
CC	This cDNA encoding for plasminogen was isolated from a lambda phage
CC	library screened with a partial cDNA clone. This cDNA is used in a new
CC	process for the production of plasminogen where a first DNA sequence
CC	encoding plasminogen and at least one additional DNA sequence encoding a
CC	protein that processes or stabilises the plasminogen is introduced into
CC	a eukaryotic host cell. The protein is selected from alpha-1-antitrypsin
CC	(AAV) and its variants and Argseirpin. Both the DNA sequences are
CC	operably linked to transcriptional promoter and terminator sequences. The
CC	host cell is cultured under conditions that allow the DNA sequences to be
CC	expressed and the recombinant plasminogen is isolated from the host cell.
CC	Co-expression of plasminogen and the protein gives increased yields of
CC	undegraded plasminogen.
XX	
SQ	Sequence 2679 BP; 751 A; 661 C; 670 G; 596 T; 1 other;
XX	
Query Match	24.5%; Score 1226.8; DB 18; Length 2679;
Best Local Similarity	80.7%; Pred. No. 0;
Matches 1503;	Conservative 1; Mismatches 278; Indels 80; Gaps 3
OY	3159 TATCGAGGCAATTCCTCCACACCACGTGCACAGGAAGACCTGTCAATCTGGTCACTTANG 3218
Dd	868 TATCGATGGGATGTGGTGTTACCTCCTGCCGGGACACCTGTCAAGCATGTGATGTCACAG 927
OY	3219 ATACCAACATGGGATCGAGGATCCAGAACCCAGAAAATCCACAATGCTGGCGTAGCCAGAAC 3278
Dd	928 ACCCTTACACACATAACAGCACACAGAAAATCTCCCCCTGCAAAAATTGGATGAAGAC 987
OY	3279 TACTGACAGAATCCAGATTCTGGGAAACAACCTGGTGTTCACACACACGATCCGTGTG 3338
Dd	988 TACTGCGCATCTCTGACGAAAAAGGGCCCATGTGCTCCATTAACACCAACAGCCAGATG 1047

QY	3339	AGG1GGAGTACTGCAATCTGACAACTGCTGAGAAACGAATCAGGTGCTTAAGACT	3339
Db	1048	CGG1GGGAGTACTTAAGATACGGTCTG-----	1076
QY	3339	CCCACTGTTGTTCCAGTTCACAGCATGGAGGCTCATCTGGAACAGCACACACTGAGCA	3458
Db	1077	-----TGACTGCTCTCCCAAGTATCCAGGAACAAATTGGCTCCCAACAGCACCACTGAGCTA	1131
QY	3459	ACCCTGTGTCGCGGAGTGTACCAATGGTAATGGCCAGAGTTATCGAGGGCATCTTCC	3518
Db	1132	ACCCTGTGTCGCGGAGTGTACCAATGGTAATGGCCAGAGTTATCGAGGGCATCTTCC	1199
QY	3519	ACCACTGTCCAGGAGAGAGATGTCAATCTTGGTCAATCCATGACACACACCGGATAG	3578
Db	1192	ACAACCAACACAGSAAAGAGTGCAGTCTGGTCAATCATGACACACACCGGACAG	1251
QY	3579	AGGACCCCAAGAAATCTACCAATGATGGCTGACAAATGACTACTGACAGAAATCCAAAT	3638
Db	1252	AAGACCCCAAGAAATCTACCAATGATGGCTGACAAATGACTACTGACAGAAATCCAAAT	1311
QY	3639	GCCGATACAGCCCTTGTTGTTTACCATGAGCCCAAGATCAGGTGGGAGTACTGACAC	3688
Db	1312	GCCGATTAAGGCCCTCGTGGTTTTACCACAGACCCCAAGGTGAGGTGGAGTACTGACAC	1371
QY	3699	CTGACGCATCTCAGACACAGAAAGGACCTGTGTGCTCTCCACTGTACTCCAGTT	3758
Db	1372	CTGAAAAATCTCAGAGAACAGAGCCAGTGTGTGACACTCCGCTGTCTCTGCTT	1431
QY	3759	CCAAGCCTTATGGCCCTCTCTGCAACAAGATGTATTTGGGAAATGGAAAGATATCCG	3818
Db	1432	CCAGATGTATGAGACTCTCTCCGAAAGAAAGTATGTATGTTGGGAAATGGAAAGATATCCG	1491
QY	3819	GCGAAGAGGCAACCACTGTTACTGGGAGCCCATGCCAGAAATGGGCTCCACAGAGCC	3878
Db	1492	GCGAAGAGGAGCACCTGTTACTGGGAGCCCATGCCAGAACTGGGCTCCACAGAGCC	1551
QY	3879	CATAGCACACACAGTTCATTCAGGGACCAAAATATGGGACAGTCTGGAAAAAAATATAC	3938
Db	1552	CATAGCACACACATTTTACCTCCAGAGAAATCCACGGGCGGGTGTGAAAAAAATATAC	1611
QY	3939	TGCCGTACCTGATGGTGATCAATCAAGTGCCTGCTTACACATGAAATCCAGAAATA	3999
Db	1612	TGCCGTACCTGATGGTGATGATGGTGCCTGCTTACACATGAAATCCAGAAATA	1671
QY	3999	CTTTTGTACTACTGTGATATCCCTCTCTGTGCATCCCTTCATTTGATTTGGGAAGCT	4058
Db	1672	CTTTTGTACTACTGTGATATCCCTCTCTGTGCATCCCTTCATTTGATTTGGGAAGCT	1731
QY	4059	CAGGTGAGCCGAGAAATGTCTCGGAAGCATTTGTAGGGGGGTGTGGCCACACCCACAT	4118
Db	1732	CAGGTGAGCCGAGAAATGTCTCGGAAGGGTGTGTAGGGGGGTGTGGCCACACCCACAT	1791
QY	4119	TCTGGGCCCTGGCAAGTCACTCTAGAAACAAGGTTTGGAAAGCACTTGTGTGAGGACAC	4178
Db	1792	TCTGGGCCCTGGCAAGTCACTCTAGAAACAAGGTTTGGAAAGCACTTGTGTGAGGACAC	1851
QY	4179	TTAATATCCCAAGATGGGTGCTGACGTGCTGACATGCTTAAAGAAAGCTCAAGGCT	4238
Db	1852	TTAATATCCCAAGATGGGTGCTGACGTGCTGACATGCTTAAAGAAAGCTCAAGGCT	1911
QY	4239	TCAATCTCAAGGTCAATCTGTGGGTGACACACCAAGAGTGAACCTGCAATCTATGTTAC	4298
Db	1912	TCAATCTCAAGGTCAATCTGTGGGTGACACACCAAGAGTGAATCTGCAACCGCATTTTAC	1971
QY	4299	GAAATAGAAAGTGTAGAGGTGTTCTTGGAGCCCAACCAAGCAAGATTTTGGCTTGTAAAG	4358
Db	1972	GAAATAGAAAGTGTAGAGGTGTTCTTGGAGCCCAACCAAGAAATTTTGGCTTGTAAAG	2031
QY	4359	CTAAGCAGGCTGCCTATCACTGACAAAGTAATCCAGCTTGTCTGCATCCCCAGAC	4418
Db	2032	CTAAGCAGTCTCTGCCTATCACTGACAAAGTAATCCAGCTTGTCTGCATCCCCAAT	2091


```
Db 1535 GGCAGAGGGGCGACCTGTTACTGGAGGCCATGCCAGAGCTGGGCTGCCAGAGCC 1594
OY 3879 CATAGACACAGCAGTTTCATTTCCAGGAGCAAAATMAATG6GCA8GTCGTGAAAAAATTAC 3938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1595 CATAGACACAGCATTTTTCACCTCCAGAGCAAAATCCAGGGGGGCTGTGAAAAAATTAC 1654
OY 3939 TGCCGTAACTGTATGTCATCATATGTCCTGGTGTCTACCAATGAATCCAGAAAA 3998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1655 TGCCGTAACTGTATGTCATCATATGTCCTGGTGTCTACCAATGAATCCAGAAAA 1714
OY 3999 CTTTGTGACTGTCATGATATCCCTCTCTGTCATCCCTTCATTTATTTGGAAGCCT 4058
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1715 CTTTGTGACTGTCATGATATCCCTCTCTGTCATCCCTTCATTTATTTGGAAGCCT 1774
OY 4059 CAAGTGAAGCCGAGAAATGTCTCGAAGCATTTGAGGGGGGTGTGTGAGCCACCCAT 4118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1775 CAAGTGAAGCCGAGAAATGTGTCTGAGGGTGTGTGAGGGGTGTGTGAGCCACCCAT 1834
OY 4119 TCCTGGCCCTGGCAAGTCAGTCAGAACAGGTTTGGAAAGCATTCCTGTGGAGGACCC 4178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1835 TCCTGGCCCTGGCAAGTCAGTCAGAACAGGTTTGGAAAGCATTCCTGTGTGGAGGACCC 1894
OY 4179 TTAATATCCCAAGAGTGGGTGCTGACTGTCCTCATCTGCTTGAAGAAGTCTCAAGGCT 4238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1895 TTGATATCCCAAGAGTGGGTGCTGACTGTCCTCATCTGCTTGAAGAAGTCTCAAGGCT 1954
OY 4239 TCATCCCTAAGTCATCCCTGGGTGCAACCAAGAGTGAACCTCAATCTCATGTTTCAG 4298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1955 TCATCCCTAAGTCATCCCTGGGTGCAACCAAGAGTGAACCTCAATCTCAATCCGATGTCAG 2014
OY 4299 GAAATGAAGTGTCTAGGCTGTTCTTGAGGCCACACACAGCATATTCCTTGTCTAAG 4358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2015 GAAATGAAGAGTGTCTAGGCTGTTCTTGAGGCCACACAGCAAGAAAGATTTGCTTGTCTAAG 2074
OY 4359 CTAAGCAGGCTGCGGTCATCTACAGCAAGTAATGTCAGCTGTGTGTCATCCCGAGAC 4418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2075 CTAAGCAGTCTGCGGTCATCTACAGCAAGTAATCCAGCTGTGTGTCATCCCGAAT 2134
OY 4419 TACATGTCACCGCAGAGCTGAATGTACATCACTGGCTGGGAGAAACCCAGGTACC 4478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2135 TATGTGTGCGCTGACCGGACCGAATGTTTATCATCACTGGCTGGGAGAAACCCAGGTACT 2194
OY 4479 TTTGGAGCTGGCTTCTCAGAGAACCCAGCTCTTGTATTTAGAAATGAATGTCAT 4538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2195 TTTGGAGCTGGCTTCTCAGAGAACCCAGCTCTTGTGTATGAATGAAGTGTCAAT 2254
OY 4539 CACTATAGT-----ATATTTGTGCTGACATTTGGCC 4571
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2255 CGCTATGAGTTTCTGAATGGAAGAGTCCATCCACGAACTCTGTGTGGGCAATTGGCC 2314
OY 4572 AGAGGCACTGACAGTGGCCAGGTCAGACAGTGAAGGCGCTGTGTTGCTTCGAGAGAC 4631
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2315 GGAGGCACTGACAGTGGCCAGGTCAGACAGTGAAGGCTCTGTGTTGCTTCGAGAGAC 2374
OY 4632 AAATACATTTTACAAGAGTCACTTCTGGGCTTGTGGTGTGTCAGCCCCATTAAGCCT 4691
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2375 AAATACATTTTACAAGAGTCACTTCTGGGCTTGTGGTGTGTCAGCCCCATTAAGCCT 2434
OY 4692 GGTGTCTATGCTGTGTTCAGGTTTGTACTTGTGATGAGGAAATGAATTAAT 4751
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2435 GGTGTCTATGCTGTGTTCAGGTTTGTACTTGTGATGAGGAAATGAATTAAT 2494
OY 4752 TAAATGAGCGGAGACAGTGAAGCATTCACTTAAGTGAAGTGAAGTGGGTGAAG 4811
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2495 TAAATGAGCGGAGACAGTGAAGCATTCACTTAAGTGAAGTGAAGTGGGTGAAG 2554
OY 4812 ATTTAGCATGCTGCAATTAATAGACAGCAATCAACAGACACTGTTCCAGACTACAG 4871
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2555 ATTTAGCATGCTGCAATTAATAGACAGCAATCAACAGACACTGTTCCAGACTACAG 2614
OY 4872 CTAATGCAAACTTGCGATTTTGTGTATTTTGTGTATTAAGCTTTTAAGTCTGATGAC 4931
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2615 CTAATGCAAACTTGCGATTTTGTGTATTTTGTGTATTTTGTGTATTTTGTGTATTT 2657
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
OY 4932 AAATCTGATTAAGTGTGATAGCTATGACATTTGTTAAAAATTAACCTGTGACTTATT 4991
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2658 GGAATTCGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2717
OY 4992 TTGATTT 4998
    || ||
Db 2718 TTGATTT 2724
```

RESULT 8
AAQ12554
ID AAQ12554 standard; DNA; 2756 BP.

XX
AC AAQ12554;
XX
DT 23-SEP-1991 (first entry)
XX
DE Encodes Plasminogen muteln T17 with thrombin cleavage site.
XX
KW protease; fibrinolysis; blood clotting; ss.

XX
FH Key Location/Qualifiers
FT mutation 179..1801
FT /tag= a
FT /note= "Ser codon (AGT) inserted"

XX
PN M09109118-A.
XX
PD 27-JUN-1991.

XX
PF 07-DEC-1990; 90WO-G001912.
XX
PR 07-DEC-1989; 89GB-0027722.
XX
PR 07-DEC-1990; 90WO-GB01911.

XX
PA (BRB1-) BRIT BIO-TECHN LTD.
XX
PI Dawson KM, Edwards RM, Forman JM;

XX
DR WPL: 1991-208145/28.
XX
DR P-PSDB; AAR12945.
XX
PT Activatable fibrinolytic and antithrombic proteins - activated by
XX
PT e.g. factor Xa, thrombin or activated protein C

XX
PS Claim 16; Fig 2 and Fig 5; 73pp; English.
XX
SS

CC This sequence encodes a plasminogen mutant cleavable by thrombin.
CC Activation is localised to the thrombus because cleavage to plasmin
CC is by an enzyme of the blood clotting pathway. Compositions
CC comprising the mutant plasminogen encoded by this sequence are used
CC for treatment or prevention of thrombosis, etc.
XX
CC See AAQ12542-Q12558.

XX
SQ Sequence 2756 BP; 766 A; 671 C; 696 G; 623 T; 0 other;

Query Match 24.4%; Score 1219.2; DB 12; Length 2756;
Best Local Similarity 80.7%; Pred. No. 0;
Matches 1509; Conservative 0; Mismatches 278; Indels 83; Gaps 4;

```
OY 3159 TATGAGGCAATATCCACACACTGTACAGAGAGAGCTGTCAATTTGTCATCTATG 3218
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 911 TATGCGGGAGATGTGGTGTATTTACCGTGTCCGGGACACCTGTACACACTGGAGTGAACG 970
OY 3219 ATACCACTGCGATCAGAGAGCCCGAGAAACTTACCCAAATGCTGCGCTGACGAGAAC 3278
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 971 ACCCTCAGACACATTAACAGAGACACCAAGAAACTTTCCTGCAAAATTTGATGTAATC 1030
OY 3279 TACTGAGGAAATCCAGATTTGTGGAAACAACCTGGTGTATACCAACGATCCGTGTGTG 3338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1031 TACTGCCCAATCTGTACGAGAAAGAGGCCCATGTGTGTCATACCAACAGCAAGTG 1090
```

Db	2135	AATTATGCTGTCGCGACCGGACCGGAATGTTTATCATCTAGCTGCTGGGAGAAACCAAGGT	2196
Qy	4476	ACCTTTGGGACGTGGCCTTCTCAGAAAGCCCAAGCTCTTTGTTATTGGAATGAAGTGTCC	4533
Db	2195	ACTTTTGGAGCGGCGCTTCTCAGAAAGCCCAAGCTCTGTTGATTGGAATGAAGTGTCC	2255
Qy	4536	AATCACTATTAAGT-----AATTTGTGCTGAGCAATTTG	4566
Db	2255	AATCGCTATGAGTTTCTGAAATGGAAGATGCCAATCCACGAACCTGTGCTGGGCAATTTG	2314
Qy	4569	GCCAGAGGACATGACACAGTTCGACAGGTCGACAGTGAAGGAGCGCTCGTTGCTTCGAGAA	4628
Db	2315	GCCGAGGACATGACAGATTTGCGAGGGTGAACAGTGGAGGTCCCTGTGGTTTCTTCGAGAA	2374
Qy	4629	GACAAATACATTTTACAAAGAGTCACTTCTTGGGGTCTTGGCTGTGCACGCCCAATAAG	4688
Db	2375	GACAAATACATTTTACAAAGAGTCACTTCTTGGGGTCTTGGCTGTGCACGCCCAATAAG	2433
Qy	4689	CCGAGTGTATGACCCGTTTTCAGAGTTTGTACTTGTGATTTAGGGAATGATGAGAAAT	4746
Db	2435	CCGAGTGTATGATTCGTTTCAAGGTTTGTACTTGTGATTTAGGGAATGATGAGAAAT	2494
Qy	4749	AATTTAATTTGACGGGAGACAGATGAGTCAACCTACTAGAGCTGAAGCTGAACGTGGGTA	4808
Db	2495	AATTTAATTTGACGGGAGACAGATGAGTCAACCTACTAGAGCTGAAGCTGGGTA	2554
Qy	4809	AGCATTTTACATGCTGCGGAAATTAATAGCAACGATCAAAAGGAAAGCACTGTTCACACTAC	4866
Db	2555	GGAATTTTACATGCGGAAATTAATGCTGCAATATCAAAAGGAAAGCACTGTTCACACTAC	2614
Qy	4869	CACCTATGCGAAACCTTGGCATTTTGTGTAATTTTGTGTATTAAGCTTTTAAGTCTGACT	4928
Db	2615	CACCTATGCGAAACCTTGGCATTTTGTGTAATTTTGTGTATTAAGCTTTTAAGTCTGACT	2657
Qy	4929	GACAATTTCTGATTAAGGTCGTCATGCTATGACATTTGTTAAATAATACCTGTGACT	4988
Db	2658	GCTGGATTTCTGTAGTAGAGGTGACATGACTATGACATTTGTTAAATAATACCTGTGACT	2711
Qy	4989	ATTTTGAATTT 4998	
Db	2718	AACCTTGATTT 2727	
RESULT 9			
AAQ12548			
ID	AAQ12548	standard; DNA; 2756 BP.	
XX	AAQ12548;		
AC			
XX	23-SEP-1991	(first entry)	
DT			
DE	Encodes Plasminogen muteln T2 with thrombin cleavage site.		
XX			
KM	protease: fibrinolysis; blood clotting; ss.		
XX			
FR	Key	Location/Qualifiers	
FT	mutation	1796..1801	
FT		/*tag= a	
FT		/note= "replaces CCTGGA (Pro-Gly) with GGTGTCA	
FT		(Gly-Gly-Pro)"	
FT	CDS	65..2500	
FT		/*tag= b	
FT		/product= modified plasminogen	
XX			
XX	W09109118-A.		
PN			
PD	27-JUN-1991.		
XX			
PE	07-DEC-1990;	90WO-G001912.	
XX			
PR	07-DEC-1989;	89GB-0027722.	
PR	07-DEC-1990;	90WO-G001911.	

XX
PA (BRI-) BRIT BIO-TECHN LTD.
XX
PI Dawson KM, Edwards RM, Forman JM;
XX
DR WPI: 1991-208145/28.
XX P-PSDB: AARI2939.
XX
PT Activable fibrinolytic and antithrombotic proteins - activated by
PT e.g. factor Xa, thrombin or activated protein C
XX
PS Claim 16; Fig 2 and Fig 5; 73pp; English.
XX
CC This sequence encodes a plasminogen mutant cleavable by thrombin.
CC Activation is localised to the thrombus because cleavage to plasmin
CC is by an enzyme of the blood clotting pathway. Compositions
CC comprising the mutant plasminogen encoded by this sequence are used
CC for treatment or prevention of thrombosis, etc.
CC See AA012542-Q12558.
XX
SQ Sequence 2756 BP; 765 A; 671 C; 697 G; 623 T; 0 other;

Query Match 24.3%; Score 1216; DB 12; Length 2756;
Best Local Similarity 80.6%; Pred. No. 0;
Matches 1507; Conservative 0; Mismatches 280; Indels 83; Gaps 4;

QY 3159 TATCGAGGATATCCCTCCACGCTGTCACAGGAAGACCTGTCATCTGTCATG 3218
DB 911 TATCGGGAATGTGGCTTTACCGTTCGGGACACCTGTCAGACTGGAGTCACAG 970
QY 3219 ATACCACTGGATCAGAGACCCAGAAAACTACCCAAATGCTGGCTGACCGAGAC 3278
DB 971 ACCCTCAGACACATTAACGAGACACAGAAAACTTCCCTGCAAAATTGGATGAAAC 1030
QY 3279 TACTGAGAAATCCAGATTCGGAAGAACACCCCTGGTGTACAGCAACGATCCGTG 3338
DB 1031 TACTGCCGAAATCTTACGGAAGAAAGGCCCATGTCCTTACACCAACAGCCAAATG 1090
QY 3339 AGGTGGAGTACTGCAATCTGACAAATGCTCAGAAACGAATGAGTGTCTAGAGACT 3398
DB 1091 CGGTGGAGTACTGTAAGTATCCGTCCTG----- 1119
QY 3399 CCACGTGTGTCACAGTTCACAGGAGGCTCATCTGTAAGGACGACCACTGAGCAA 3458
DB 1120 -----TGACTCCCTCCCACTATCCAGGAACATTTGGCTCCACAGCACCTGAGCTA 1174
QY 3459 ACCCTGTGTGTCGGGACGTCACCATGTATGTCGACAGATTATCGAGGACATTTCTCC 3518
DB 1175 ACCCTGTGTGTCGGGACGTCACCATGTATGTCGACAGATTATCGAGGACATTTCTCC 1234
QY 3519 ACCAGTGCACAGGAAGATGCAATCTGTCATGTCATGCAACACACACCCGACATCAG 3578
DB 1235 ACCACACACAGGAAGATGTCAGTCTGTGTCATGTCATGCAACACACCCGACATCAG 1294
QY 3579 AGGACCCAGAAAACTACCAATGATGCGCTGACATGAATCTACTGAGGAATCCAGAT 3638
DB 1295 AAGACCCAGAAAACTACCAATGATGCGCTGACATGAATCTACTGAGGAATCCAGAT 1354
QY 3639 GCGGATACAGGCGCTGTGTTTACATGAGGCCACACATCAGTGGGAGTACTGCAAC 3698
DB 1355 GCGGATACAGGCGCTGTGTTTACATGAGGCCACACATCAGTGGGAGTACTGCAAC 1414
QY 3699 CTGACGATGTCACAGACAGAGGATGTGTGCTCTCCGATGTCATCAGATG 3758
DB 1415 CTGAAAAAATGCTCAGGACAGAGGATGTGTGAGACCTCCGCTGTTCTCTCTT 1474
QY 3759 CCAAGGCTAGGCGCTCTCTCTGACAGACATGTATGTTGGATGGGAAGATACCGG 3818
DB 1475 CCAAGATGTAGACCTCTCTCTGAAAGACATGTATGTTGGGAATGGAAAGGATACCGG 1534
QY 3819 GGCAGAGGACCACTGTTACTGAGAGCCATGCCAGATGGGCTGCCAGAGACCC 3878
DB 1535 GGCAGAGGCGCACACTTACTGAGAGCCATGCCAGAGACTGGGCTGCCAGAGACCC 1594

QY 3879 CATAGACAGACAGCTTCATTCACGAGGACAAATTAATGGGCGAGTCTGGAATAATTTAC 3938
DB 1595 CATAGACAGACAGCTTCATTCACGAGGACAAATTCACGAGGCGGCTGTGAAAAAATTAC 1654
QY 3939 TGCCGTAAACCTGATGTCATCAATGATGTCCTGGTGTACACAAATGATCAAGAAAA 3998
DB 1655 TGCCGTAAACCTGATGTCATGATGTCCTGGTGTACACAAATGATCAAGAAAA 1714
QY 3999 CTTTGTACATGTCATGATGTCCTGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 4058
DB 1715 CTTTGTACATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 1774
QY 4059 CAACTGAGCCGGAAGAAATGCTCTG--AAGCATGTAGAGGGGCTGTGGCCACCA 4115
DB 1775 CAACTGAGCCGGAAGAAATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 1834
QY 4116 CATTCCTGGCCGTCGAGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 4175
DB 1835 CATTCCTGGCCGTCGAGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 1894
QY 4176 ACCTTAATATCCGAGAGTGGTGTGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 4235
DB 1895 ACCTTAATATCCGAGAGTGGTGTGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 1954
QY 4236 CCTTCATCCTCAAGAGTCATCCTGGGTCACACAAAGTGAACCTGGAATGTCATGTT 4295
DB 1955 CCTTCATCCTCAAGAGTCATCCTGGGTCACACAAAGTGAACCTGGAATGTCATGTT 2014
QY 4296 CAGGAATAGAAATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 4355
DB 2015 CAGGAATAGAAATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 2074
QY 4356 AAGCTAAGCAAGGCTGCTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 4415
DB 2075 AAGCTAAGCAAGGCTGCTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 2134
QY 4416 GACTACATGTCACGCGCAAGTGAATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 4475
DB 2135 AATTAATGTCGTCGTCACCGCAAGTGAATGTCATGTCATGTCATGTCATGTCATGTCATG 2194
QY 4476 ACCTTGGAGTGGCTTTCAGAGGAAGCCAGCTCTGTCATGTCATGTCATGTCATGTCATG 4535
DB 2195 ACCTTGGAGTGGCTTTCAGAGGAAGCCAGCTCTGTCATGTCATGTCATGTCATGTCATG 2254
QY 4536 AATCATATAAGT-----AATTTGTGTCGAGCATTTG 4568
DB 2255 AATCATATAAGT-----AATTTGTGTCGAGCATTTG 2314
QY 4569 GCGAGAGGCACTGACATGTCAGAGGTCAGAGGTCAGAGGTCAGAGGTCAGAGGTCAGAGG 4628
DB 2315 GCGAGAGGCACTGACATGTCAGAGGTCAGAGGTCAGAGGTCAGAGGTCAGAGGTCAGAGG 2374
QY 4629 GACAAATACATTTTACAAGAGTCACTTCTGAGGTCCTGAGTGTGCAAGCCCAATAG 4688
DB 2375 GACAAATACATTTTACAAGAGTCACTTCTGAGGTCCTGAGTGTGCAAGCCCAATAG 2434
QY 4689 CCGTGTGTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 4748
DB 2435 CCGTGTGTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2494
QY 4749 AATTAATGAGCGGAGACAGAGTGAAGCATCACTACTTGAAGCTGAAACGTGGGTA 4808
DB 2495 AATTAATGAGCGGAGACAGAGTGAAGCATCACTACTTGAAGCTGAAACGTGGGTA 2554
QY 4809 AGGATTAATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 4868
DB 2555 AGGATTAATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 2614
QY 4869 CAGCTATGCCAAACCTTGGCATTTTGTATTTGTATTAAGCTTTAAGGTCGTCAGT 4928
DB 2615 CAGCTATGCCAAACCTTGGCATTTTGTATTTGTATTAAGCTTTAAGGTCGTCAGT 2657

Db	1031	TACTGCCGCAATCCTGACGGAAAAAGGCCCATGCTGGCCATACCAACCAAGCCAAATG	1090
Qy	3339	AGGTGGAGATACAGCAATCTGACACAAATGCTCAGAAACAGAAATCAGTGTCTAGAGACT	3398
Db	1091	CGGTGGGAGTACTGTAAATACCGTCTCTG-----	1119
Qy	3399	CCCACTGTGTCCAGTTCACAGCATGAGGGCTCATTTCTGAAGCAGCAACCAATGAGCA	3458
Db	1120	-----TGACTCTCTCCCAAGTATTCACGGAACAATATGGTCTCCACAGCAACCTGAGCTA	1174
Qy	3459	ACCCCTTGGTTCGGGAGTGTCTACCATGTATATGGCCAGAGTTATCCAGGCACATTTCTCC	3518
Db	1175	ACCCCTTGGTTCAGGAGCTGTCTACCATGTGTGATGGACAGAGTACCGGAGCAATCTCTCC	1234
Qy	3519	ACCACTGTCAAGAAAGAGCATGTCTCAATCTTGGTCTCCATGATCAGACCAACACCGGATCAT	3578
Db	1235	ACCAACCAACCAAGAAAGAGTTCAGTCTTGCTTGTCTATATACCAACCAACCGGACACG	1294
Qy	3579	AGGACCCCAAGAAACATACCCAAATGATGGCCAGACATGAACTGACAGGATTCAGAT	3638
Db	1295	AAGACCCCAAGAAACATACCCAAATGCTGGCCAGACATGAACTGACAGGATTCAGAT	1354
Qy	3639	GCGCATACAGGCCCTTGTGTATTTACCATGACCCAGCATAGGTGGAGTACTGCAAC	3688
Db	1355	GCCCATAAAGGCCCTGTGTGTATTTACCAAGACCCAGCGTCGTGGAGTACTGCAAC	1414
Qy	3699	CTGACGGATGTCTACAGACAGAAAGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3758
Db	1415	CTGAAAAAATGTCTCAGGACAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1474
Qy	3759	CCAAGCCTTGAAGGCTCTCTCTGTGAACAAGACTGATGTGTGTGTGTGTGTGTGTGTGT	3818
Db	1475	CCAAGTGTAGAGACTCTCTCTCGAAGAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT	1534
Qy	3819	GCGAAGAAGCAACACTGTTACTGGAGGCCATGCGAAGAAATGGGCTGTGCCAGAGCC	3878
Db	1535	GCGAAGAAGGCGACACTGTTACTGGAGGCCATGCGAAGAAATGGGCTGTGCCAGAGCC	1594
Qy	3879	CATGACACACAGACGTTCAATCCAGGACAAATTAATGGGAGGTCTGGAAAAAATATAC	3938
Db	1595	CATGACACACAGACGTTCAATCCAGGACAAATTAATGGGAGGTCTGGAAAAAATATAC	1654
Qy	3939	TGCCGTAAACCTGATGGTGCATCATGTGCTCCGTGTGTGTGTGTGTGTGTGTGTGTGT	3998
Db	1655	TGCCGTAAACCTGATGGT	1714
Qy	3999	CTTTTGTACTGTGTATCCCTCTCTGTGTGCATCTCTTCAATTTGATGTGGAAAGCT	4058
Db	1715	CTTTTGTACTGTGTATCCCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1774
Qy	4059	CAATGGAGCGGAAGAAATCTCT-----GGAAGCATTTGTAGGGGGTGTGTGCC	4109
Db	1775	CAATGGAGCGGAAGAAATTTGTGTATCATAGACGGAAGATTTGTGGGGGGTGTGTGCC	1834
Qy	4110	CACCCATTTCTGGCCCTGGCAAGCAGTCAACAGCAAGTTTGGAAAGGACATCTGT	4159
Db	1835	CACCCATTTCTGGCCCTGGCAAGCAGTCTTGAACAAGGTTTGGATGTGACTCTGT	1894
Qy	4170	GGAGGACATTAATATCCAGAGTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	4229
Db	1895	GGAGGACATTTGATATCCCAAGTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1954
Qy	4230	TCAAGGCTTCTATCTCTAAGAGTCAATCTGTGGTGCACACCAAGAAAGTGAACCTGATCT	4289
Db	1955	CCAAGGCTTCTATCTCTAAGAGTCAATCTGTGGTGCACACCAAGAAAGTGAATCTGCAACG	2014
Qy	4290	CATGTTCAAGAAATAGAAGTGTGTAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	4349
Db	2015	CATGTTCAAGAAATAGAAGTGTGTAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2074
Qy	4350	TTTGCTAAGCTAAGCAGGCTGTGCGTATCACTGACAAAGTAATGCCAGTTGTCTGCA	4409

Db	2075	TTGGTAAAGCTAAGCAAGTCTCTCCCGTCATCCTGACAGCAAAAGTAATCCACAGCTTGTGTGCCA	2134
QY	4410	TCGCCAGCTACATAGTGTCCACCGCAGGAGCTGAATGTTTACATCATGGCTGGGGAGAAAC	4469
Db	2135	TCGCCAATTAATGTGTGTGCTGTGACCGGACCGAATGTTTATCATCTGGCGTGGGAGAAAC	2194
QY	4470	CAAGGTACTTTGGGACTGGCCCTTCTCAGAGAAAGCCAGCTCCTGTTATTGACATATAA	4529
Db	2195	CAAGGTACTTTTGGAGCTGGCCCTTCTCAGAGAAAGCCAGCTCCTGTTATTGAGATATAA	2254
QY	4530	GTGGCATCATCCTTAAGT-----ATATTGTGTCTAG	4562
Db	2255	GTGTGCATTCGTATGAGTGTTCGAATGAGAGTCCAAATCCACCGAAGCTGTGTCTGG	2314
QY	4563	CATTGGCAGAGGACATGTGACAGTGGCAGGGGTGACAGTGAAGGGGCTCTGGTTGCTTC	4622
Db	2315	CATTGGCGCGAGGACATGTGACAGTGGCAGGGGTGACAGTGAAGGGGCTCTGGTTGCTTC	2374
QY	4623	GAGAAAGCAAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTGTGGCTGTACGCCCC	4682
Db	2375	GAGAAAGCAAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTGTGGCTGTACGCCCC	2434
QY	4683	AATAAGCCTGGTGTCTATGCTCGTGTTCAGAGTTTGTACTTGTGATGAGGAGATATGATG	4742
Db	2435	AATAAGCCTGGTGTCTATGCTCGTGTTCAGAGTTTGTACTTGTGATGAGGAGATATGATG	2494
QY	4743	AGAAATATTAATTGAGCGGAGACAGAGTGAAGCATCAACCTACTTAGAGCTGAACG	4802
Db	2495	AGAAATATTAATTGAGCGGAGAGACAGAGTGAAGCATCAACCTACTTAGAGCTGAACG	2556
QY	4803	TGGGTAAAGATTTAGCATGCTGTGAAATATATAGACAAATCAAGACAGACACTGTGCC	4863
Db	2555	TGGGTAAAGATTTAGCATGCTGTGAAATATATAGACAAATCAAGACAGACACTGTGCC	2614
QY	4863	AGCTACCAAGCTATGCGCAACCTTGGCATTTTGTGTATTTTGTGTATTAAGT	4922
Db	2615	AGCTACCAAGCTATGCGCGAAACCTGGCATTTTGTGTATTTTGTGTATTT-----TT	2657
QY	4923	CTGACTGACAAATCTGTATTAAGGTGTACATGCTATGACATTTGTTAAATAATACCT	4982
Db	2658	CTGACTGCTGATTTCTGTAGTAAGGTGTACATGCTATGACATTTGTTAAATAATACCT	2717
QY	4983	GCACCTATTTTGATTT	4998
Db	2718	GTACTTAACCTTTGATTT	2733
RESULT 13			
AAQ12543			
ID	AAQ12543	standard; DNA; 2759 BP.	
AC	AAQ12543;		
XX	23-SEP-1991	(first entry)	
DE	Encodes Plasminogen mutlein X2 with factor Xa cleavage site.		
XX	protease; fibrinolysis; blood clotting; ss.		
KM			
PH	Key	Location/Qualifiers	
FT	mutation	1796..1804	
FT		/*tag= a	
FT	CDS	65..2503	
FT		/*tag= b	
FT		/product= modified plasminogen	
PN	W09109118-A.		
PD	27-JUN-1991.		
PF	07-DEC-1990;	90WO-G001912.	
XX			


```

Db 2075 TTGGTAAAGCTAAGCAGTCCGCGGTATCATCTAGACAAAGTAATCCAGCTTGTCTGCA 2134
OY 4410 TCCCGAGCTACATGCTACACGGCCAGAGCTAATTAATCATCATGCTGGGGAGAAC 4469
Db 2135 TCCCAAAATTTATGTGTCGCTGACCGGACGAAATTTTCTCATCTGCTGGGAGAAC 2194
OY 4470 CAAGTACCTTTGGGACTGGCCTTCTCAAGGAAGCCAGCTCTTGTATGAGAAATGAA 4529
Db 2195 CAAGTACTTTTGGAGCTGGCCTTCTCAAGGAAGCCAGCTCCCTGTGATGAAATGAA 2254
OY 4530 GTGTGCATCTACTTAAGT-----ATATTGTGCTGAG 4562
Db 2255 GTGTGCATCTGCTATGAGTTCTGAATGGAAGAGTCCAAATCCACCGAACTCTGTGCTGGG 2314
OY 4563 CATTTGGCCAGAGCAGTACAGTATGCGAGGAGTACAGTGGAGGGCCCTGGTTGTTC 4622
Db 2315 CATTTGGCCAGAGCAGTACAGTATGCGAGGAGTACAGTGGAGGGCCCTGTGTTGTTC 2374
OY 4623 GAGAAGGACAAATACATTTTACAAGAGTCACTTTGGGGTCTTGGCTGTGCAGGCC 4682
Db 2375 GAGAAGGACAAATACATTTTACAAGAGTCACTTTGGGGTCTTGGCTGTGCAGGCC 2434
OY 4683 AATTAAGCCTGTCTATGCTCGTGTTCAGAGTTTGTACTTGTGATTGAGGAAATG 4742
Db 2435 AATTAAGCCTGTCTATGCTCGTGTTCAGAGTTTGTACTTGTGATTGAGGAAATG 2494
OY 4743 AGAATTAATTAATTTGAGCGGAGAGAGTGAAGTCAATCAACCTTCTTAAGAGGAGAAC 4802
Db 2495 AGAATTAATTAATTTGAGCGGAGAGAGTGAAGTCAATCAACCTTCTTAAGAGGAGAAC 2554
OY 4803 TGGGTAAGATTTAGCATGCTGTAATTAATTAAGAGCAATCAAGAGACACTGTTC 4862
Db 2555 TGGGTAAGATTTAGCATGCTGTAATTAATTAAGAGCAATCAAGAGACACTGTTC 2614
OY 4863 AGCTACAGCATATGCCAAACCTTGGCATTTTGTGATTTTGTGTAATTAAGT 4922
Db 2615 AGCTACAGCATATGCCAAACCTTGGCATTTTGTGATTTTGTGTAATTAAGT 2657
OY 4923 CTGACAGCAAAATCTGATTAATTAAGTGTATAGTATGATTAATTAATTAAGT 4982
Db 2658 CTGACAGCATATGATTAATTAAGTGTATAGTATGATTAATTAATTAAGT 2717
OY 4983 GCACCTATTTGATTT 4998
Db 2718 GTACTTAACCTTGAAT 2733

RESULT 15
AAQ12545
ID AAQ12545 standard; DNA: 2762 BP.
AC AAQ12545;
XX
DT 23-SEP-1991 (first entry)
DE Encodes Plasminogen mutuin X5 with factor Xa cleavage site.
KM
XX
XX protease; fibrinolysis; blood clotting; ss.
FH
XX Key Location/Qualifiers
FT 1796..1807
FT mutation
FT /tag= a
FT /note= "Replaces CDT (Pro codon) with four codons"
FT CDS 65..2509
FT /tag= b
FT /product= modified plasminogen
XX
XX WO9109118-A.
XX
XX 27-JUN-1991.
XX
XX 07-DEC-1990; 90WO-G001912.
```

```

XX
PR 07-DEC-1989; 89GB-0027722.
PR 07-DEC-1990; 90WO-GB01911.
XX
XX (BRB1-) BRIT BIO-TECHN LTD.
XX
XX Dawson KM, Edwards RM, Forman JM;
XX
XX MPI: 1991-208145/28.
XX P-PSDB; MAR12936.
XX
XX Activatable fibrinolytic and antithrombic proteins - activated by
XX e.g. factor Xa, thrombin or activated protein C
XX
XX Claim 16; Fig 2 and Fig 4; 73pp; English.
XX
XX This sequence encodes a plasminogen mutant cleavable by Factor Xa.
XX Activation is localised to the thrombus because cleavage to plasmin
XX is by an enzyme of the blood clotting pathway. Compositions
XX comprising the mutant plasminogen encoded by this sequence are used
XX for treatment or prevention of thrombosis, etc.
XX See AAQ12542-Q12558.
XX
SQ Sequence 2762 BP; 769 A; 671 C; 698 G; 624 T; 0 other;
XX
Query Match 24.2%; Score 1210; DB 12; Length 2762;
Best Local Similarity 80.3%; Pred. No. 0;
Matches 1507; Conservative 0; Mismatches 280; Indels 89; Gaps 4;
OY 3159 TATCGAGCATATCTCCACCACTGTACAGAGAGACCTGTCAATCTTGTATCTATG 3218
Db 911 TATCGAGGGAATGTGCTGTATACCTGTCCGGACACCTGTACAGCATGTGACAG 970
OY 3219 ATTCACACTGGATACAGAGACCCAGAAATCAACCAATGCTGGCTGACCGAGAAC 3278
Db 971 ACCCTTCACACATATACAGAGACACCAAACTTCCCTGCAAAATTTGGATGAAAC 1030
OY 3279 TACTCGAGAAATCCAGATTTCTGGGAAACAACCTGTGTATACAAACGATCCGTGTG 3338
Db 1031 TACTCGCGCAATCTTACAGGAAAGAGGCCCATGTGTGATACCAACCAAGCAAGTG 1090
OY 3339 AGGTGGAGTACTGCATCTGACACATGCTCAGAAACGAATCAGTGTCTTAGAGACT 3398
Db 1091 CGGTGGAGTACTGTAATGATACCGTCTG----- 1119
OY 3399 CCACATGTTTCCATTCCAGATGAGAGGCTCATTTGAGAGCAGCAACCACTAGCAA 3458
Db 1120 -----TGACTCCTCCCAAGTATCCAGGAAACAATTTGGCTCCACAGACACCACTAGCTA 1174
OY 3459 ACCCTGTGTGTCGGCAGTACCTACCATGTTATGGCCAGAGTTATGAGGACATTTCTCC 3518
Db 1175 ACCCTGTGTGTCGGCAGTACCTACCATGTTATGGCCAGAGTACCGAGGACATCTCTCC 1234
OY 3519 ACCACTGTACAGAGAGACATGCAATCTTGTGATCCATGACACACACCGGATCAG 3578
Db 1235 ACCACACACAGAGAAAGTGTGATCTTGTGATCATATGACACACACCGGATCAG 1294
OY 3579 AGGACCCCAAGAAATCAACCAATGATGGCCCTGGAATGAATCTAGAGAAATCCAAAT 3638
Db 1295 AAGACCCCAAGAAATCAACCAATGATGGCCCTGGAATGAATCTAGAGAAATCCAAAT 1354
OY 3639 GCCGATACAGGCCCTTGGTGTATTTACATGAGACCCACAGATGAGTGGAGTACTGCAAC 3698
Db 1355 GCCGATTAAGGCCCTTGGTGTATTTACATGAGACCCACAGATGAGTGGAGTACTGCAAC 1414
OY 3699 CTGACGCGATGCTCAGACACAGAGAGAGTGTGTGCTCTCCGACATGTCATCAGGTT 3758
Db 1415 CTGAAAAAATGCTCAGAGAAAGAGAGAGTGTGTGACCATCCCGCTGTGTCTGCTT 1474
OY 3759 CCAAGCCTAGGGCCCTTCTTGAACAAGATCTATGTTGGGAATGGAAAGATACCGG 3818
Db 1475 CCAAGTGTAGAGATCTTCCAGAAAGAGACTGTATGTTGGGAATGGAAAGATACCGA 1534
```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 06:19:18 ; Search time 148 Seconds
(without alignments)
10360.701 Million cell updates/sec

Title: US-09-923-515-3_COPY_8939_13938

Perfect score: 5000

Sequence: 1 ctccgactgtaccgccgtt.....ctgcattatttgatttga 5000

Scoring table:

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_NA.*
2: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCRTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1237	24.7	2753	1 US-07-854-603-1	Sequence 1, Appli
2	1236.8	24.7	2679	6 5200340-7	Patent No. 5200340
3	1110.6	22.2	2296	1 US-07-750-080A-18	Sequence 18, Appl
4	1110.6	22.2	2296	4 US-08-651-472-18	Sequence 18, Appl
5	1110.6	22.2	2296	4 US-08-358-928-18	Sequence 18, Appl
6	1104.8	22.1	2497	1 US-08-643-219-12	Sequence 12, Appl
7	1104.8	22.1	2497	2 US-09-131-995-12	Sequence 12, Appl
8	1104.8	22.1	2497	2 US-08-832-087B-12	Sequence 12, Appl
9	1104.8	22.1	2497	3 US-08-851-350-12	Sequence 12, Appl
10	1104.8	22.1	2497	4 US-09-133-154-12	Sequence 12, Appl
11	697.6	14.0	1724	6 5200340-5	Patent No. 5200340
12	557.8	11.2	690	1 US-08-379-621-1	Sequence 1, Appli
13	557.8	11.2	690	2 US-08-889-078-1	Sequence 1, Appli
14	257.2	5.1	1134	4 US-09-206-059-29	Sequence 29, Appl
15	90.8	1.8	2172	4 US-08-030-410-2	Sequence 2, Appli
16	89.2	1.8	2184	1 US-07-815-333A-1	Sequence 1, Appli
17	89.2	1.8	2288	1 US-08-290-937B-4	Sequence 1, Appli
18	89.2	1.8	2288	1 US-07-838-410-2	Sequence 4, Appli
19	85.4	1.7	2219	1 US-07-882-925A-1	Sequence 1, Appli
20	85.4	1.7	2219	1 US-07-882-925A-2	Sequence 2, Appli
21	85.4	1.7	2219	1 US-08-184-012C-1	Sequence 1, Appli
22	85.4	1.7	2219	1 US-08-184-012C-2	Sequence 2, Appli
23	84.2	1.7	2216	2 US-08-666-082B-2	Sequence 2, Appli
24	84.2	1.7	2232	1 US-08-334-177-1	Sequence 1, Appli
25	84.2	1.7	2232	5 PCN-US95-13830-1	Sequence 1, Appli
26	84.2	1.7	2262	1 US-07-882-925A-7	Sequence 7, Appli
27	84.2	1.7	2262	1 US-08-184-012C-7	Sequence 7, Appli

28	80.4	1.6	645	3 US-08-985-526-22	Sequence 22, Appl
29	80.4	1.6	1284	3 US-08-985-526-24	Sequence 24, Appl
30	80	1.6	2188	1 US-07-882-925A-4	Sequence 4, Appli
31	80	1.6	2188	1 US-08-184-012C-4	Sequence 4, Appli
32	68	1.4	2021	1 US-07-882-925A-3	Sequence 3, Appli
33	68	1.4	2021	1 US-08-184-012C-3	Sequence 3, Appli
34	64.6	1.3	2544	4 US-09-518-046-3	Sequence 1, Appli
35	63	1.3	2413	4 US-09-518-046-1	Sequence 1, Appli
36	62.2	1.2	1130	4 US-09-386-653A-8	Sequence 8, Appli
37	62	1.2	654	4 US-09-280-116-57	Sequence 57, Appl
38	62	1.2	1110	4 US-09-386-653A-1	Sequence 1, Appli
39	60.8	1.2	6100	1 US-07-882-925A-6	Sequence 6, Appli
40	60.8	1.2	6100	1 US-08-184-012C-6	Sequence 6, Appli
41	60.6	1.2	2038	4 US-09-008-271A-18	Sequence 18, Appl
42	60.6	1.2	2079	4 US-09-656-002-1	Sequence 1, Appli
43	60.4	1.2	1109	4 US-09-088-651-1	Sequence 1, Appli
44	60.2	1.2	2416	4 US-09-261-416-1	Sequence 1, Appli
45	59.4	1.2	3398	5 PCT-US95-08493-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-07-854-603-1
: Sequence 1, Application US/07854603
: Patent No. 5637492
: GENERAL INFORMATION:
: APPLICANT: Dawson, Keith M
: APPLICANT: Edwards, Richard M
: APPLICANT: Forman, Joan M
: TITLE OF INVENTION: Activatable fibrinolytic and
: TITLE OF INVENTION: anti-thrombotic proteins
: NUMBER OF SEQUENCES: 40
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dr. John J. McDonnell
: STREET: Ten South Wacker Drive, Suite 3000
: CITY: Chicago
: STATE: IL
: COUNTRY: USA
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/854, 603
: FILING DATE: 19901207
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: McDonnell, John J
: REGISTRATION NUMBER: 26,949
: REFERENCE/DOCKET NUMBER: 92,338
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-715-1000
: TELEFAX: 312-715-1234
: TELEX: 910-221-5317
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2753 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: 1..2753
: OTHER INFORMATION: /note="Fig. 2 Plasminogen cDNA
: OTHER INFORMATION: sequence"
: FEATURE:
: NAME/KEY: sig.peptide
: LOCATION: 65..121

Query Match	Best Local Similarity	Score	DB 1	Length	2753;
Matches 1512;	Conservative	0;	Mismatches 275;	Indels 80;	Gaps 3;
US-07-854-603-1					
NAME/KEY: mat_peptide					
LOCATION: 122..2494					
FEATURE:					
NAME/KEY: CDS					
LOCATION: 65..2494					
FEATURE:					
NAME/KEY: misc_feature					
LOCATION: 54..55					
OTHER INFORMATION: /note= "Ball site"					
NAME/KEY: misc_feature					
LOCATION: 2564..2565					
OTHER INFORMATION: /note= "Spli site"					
Query Match	24.7%;	Score 1237;	DB 1;	Length 2753;	
Best Local Similarity	81.0%;	Pred. No. 0;			
Matches 1512;	Conservative	0;	Mismatches 275;	Indels 80;	Gaps 3;
US-07-854-603-1					
NAME/KEY: mat_peptide					
LOCATION: 122..2494					
FEATURE:					
NAME/KEY: CDS					
LOCATION: 65..2494					
FEATURE:					
NAME/KEY: misc_feature					
LOCATION: 54..55					
OTHER INFORMATION: /note= "Ball site"					
NAME/KEY: misc_feature					
LOCATION: 2564..2565					
OTHER INFORMATION: /note= "Spli site"					
Query Match	24.7%;	Score 1237;	DB 1;	Length 2753;	
Best Local Similarity	81.0%;	Pred. No. 0;			
Matches 1512;	Conservative	0;	Mismatches 275;	Indels 80;	Gaps 3;
US-07-854-603-1					
NAME/KEY: mat_peptide					
LOCATION: 122..2494					
FEATURE:					
NAME/KEY: CDS					
LOCATION: 65..2494					
FEATURE:					
NAME/KEY: misc_feature					
LOCATION: 54..55					
OTHER INFORMATION: /note= "Ball site"					
NAME/KEY: misc_feature					
LOCATION: 2564..2565					
OTHER INFORMATION: /note= "Spli site"					
Query Match	24.7%;	Score 1237;	DB 1;	Length 2753;	
Best Local Similarity	81.0%;	Pred. No. 0;			
Matches 1512;	Conservative	0;	Mismatches 275;	Indels 80;	Gaps 3;
US-07-854-603-1					
NAME/KEY: mat_peptide					
LOCATION: 122..2494					
FEATURE:					
NAME/KEY: CDS					
LOCATION: 65..2494					
FEATURE:					
NAME/KEY: misc_feature					
LOCATION: 54..55					
OTHER INFORMATION: /note= "Ball site"					
NAME/KEY: misc_feature					
LOCATION: 2564..2565					
OTHER INFORMATION: /note= "Spli site"					
Query Match	24.7%;	Score 1237;	DB 1;	Length 2753;	
Best Local Similarity	81.0%;	Pred. No. 0;			
Matches 1512;	Conservative	0;	Mismatches 275;	Indels 80;	Gaps 3;
US-07-854-603-1					
NAME/KEY: mat_peptide					
LOCATION: 122..2494					
FEATURE:					
NAME/KEY: CDS					
LOCATION: 65..2494					
FEATURE:					
NAME/KEY: misc_feature					
LOCATION: 54..55					
OTHER INFORMATION: /note= "Ball site"					
NAME/KEY: misc_feature					
LOCATION: 2564..2565					
OTHER INFORMATION: /note= "Spli site"					
Query Match	24.7%;	Score 1237;	DB 1;	Length 2753;	
Best Local Similarity	81.0%;	Pred. No. 0;			
Matches 1512;	Conservative	0;	Mismatches 275;	Indels 80;	Gaps 3;
US-07-854-603-1					

Db	1655	TCGCCGTAACCCCTGATGGTGAATGTAGTGTGTGCTCCGTGGTGGCTACACGACAAATTCAGAAATA	1714
Qy	3999	CTTTTGTGACTACTGTGATATCCCTCTCTGTGACATCCCTCTTTCATATTTGATTTGGAAAGCT	4058
Db	1715	CTTTCAGACTACTGTGATGTGCTCCACTGTGTGGGCCCTTCATTTGATTTGTGGAAAGCT	1774
Qy	4055	CAAGTGGAGCCGGAAGAAATGTCTCTGGAGACATTTGTATGGGGGGTGTGTGGCCACACCCACAT	4118
Db	1775	CAAGTGGAGCCGGAAGAAATGTCTCTGGAGAGGGTGTATGGGGGGTGTGTGGCCACACCCACAT	1834
Qy	4119	TCCTGGCCCTGGCAAGTCACTCTAGAAACAAGTTTGGAAGACACTTCTGTGGAGGCAC	4178
Db	1835	TCCTGGCCCTGGCAAGTCACTCTAGAAACAAGTTTGGAATGCATCTGTGTGGAGGCAC	1894
Qy	4179	TTAATATCCCAAGAGTGGGTGTGACTGTCTGCTACTGCTTGAAGAAGTCTCAAGCCT	4238
Db	1895	TTGAATATCCCAAGAGTGGGTGTGACTGTCTGCTACTGCTTGGAGAAAGTCCCAAGGCCT	1954
Qy	4239	TCATCTCAAGAGTCACTCTGCGGGGACACCAAGAAAGTGAACCTGTGAATCTCATATTCAG	4298
Db	1955	TCATCTCAAGAGTCACTCTGCGGGGACACCAAGAAAGTGAATCTGGAACCGATTTTAC	2014
Qy	4299	GAAATAGAAGTGTCTAGAGCTGTCTTGAGCCACACAAGCAAGATATTCCTTGTAAAG	4358
Db	2015	GAAATAGAAGTGTCTAGAGCTGTCTTGAGCCACACAAGAAAGATATTCCTTGTAAAG	2074
Qy	4359	CTAAGCAGAGCCTGCGCGTCACTACAGCAAAAGTAAAGTCAGCTGTGTGCAATCCCAAC	4418
Db	2075	CTAAGCAGTCTCGCGTCACTACAGCAAAAGTAAATCCAGCTTGTGTGCAATCCCAAT	2134
Qy	4419	TACATGTGCACCGCCAGAGACTGAATGTACATCACTGTGGTGGGAGAAAACCAAGTATCC	4478
Db	2135	TATGTGTGCTGTGACCGGACCGAATGTTTCATCACTGTGGGAGAAAACCAAGTATCC	2194
Qy	4479	TTTGGGACTGGCCCTTTCCAAGGAAGCCACATCCTGTTATTTAGAAATGAAGTGTGCAT	4538
Db	2195	TTTGGAGCTGGCCCTTTCCAAGGAAGCCACATCCTGTGTATTTAGAAATGAAGTGTGCAT	2254
Qy	4539	CACATATAAGT-----ATATTTGTGCTGAGCATTTGGCC	4571
Db	2255	CGCTGTAGTGTTCGTAGTGAAGAGTCCATCCACCGAATCCTGTGTGGGCAATTTGGCC	2314
Qy	4572	AGAGGCACTGACATTTTCCAGGGGTGACAGTGTGAGGGCCCTGTGTTGCTTGCAGAAAGAC	4631
Db	2315	GGAGGCACTGACATTTTCCAGGGGTGACAGTGTGAGGGGTCTGTGTTGCTTGCAGAAAGAC	2374
Qy	4632	AAATCATTTTACAGAGAGTCACTTCTTGGGGCTTGGCTGTGCACGGCCCAATTAAGCT	4691
Db	2375	AAATCATTTTACAGAGAGTCACTTCTTGGGGCTTGGCTGTGTGCACGGCCCAATTAAGCT	2434
Qy	4692	GGGTGTACTGTCTGTTCACAAGTTTGTACTTTGATTTGAGGAGTGTAGGAATAT	4751
Db	2435	GGGTGTACTGTCTGTTCACAAGTTTGTACTTTGATTTGAGGAGTGTAGGAATAT	2494
Qy	4752	TAAATGGACGGGAGACAGAGGGAAGCACTAACCTTATTTGAAGAGTGAAGACGTGGGTAAAG	4811
Db	2495	TAAATGGACGGGAGACAGAGTGTACACCTGACACTTGAAGGCTGTGAACGTGGGTAAAG	2554
Qy	4812	ATTTAGCATGCTGGAATTAATATAGACCAATCAAGGAAGACACTGTTCCAGCTACAG	4871
Db	2555	ATTTAGCATGCTGGAATTAATCTGACAGTAAATCAAGGAAGACACTGTGCCAGCTACAG	2614
Qy	4872	CTATGGCAAAACCTTGGAATTTTGGATTTTGTGTATTAAGCTTTAAGGTGTACATGAC	4931
Db	2615	CTATGGCAAAACCTTGGAATTTTGTGTATTT-----TTCTGACGTCT	2657
Qy	4932	AAATTCGTATTAAGTGTATAGCTATAGCTATGACATTTGTTAAAAATTAACCTGTGCACCTATT	4991
Db	2658	GGATTCGTATTAAGTGTATAGCTATGACATTTGTTAAAAATTAACCTGTGTACTTAAC	2717
Qy	4992	TTTGAATTT	4998

3759 CCAAGCCTAAGGCGCTCTCTCTGAACAAGACGTATGTTGGGAATGGGAAGATACCGG 3818
1200 CCAGATGTAGAGACTCTCTCCGAAGAAGACTGATGTTGGGAATGGGAAGATACCGA 1259
3819 GGCAGAGAGGCAACCTCTTACTGGAGCCATGCCAGATGGGCTCCAGAGAGCC 3878
1260 GGCAGAGAGGCAACCTCTTACTGGAGCCATGCCAGATGGGCTCCAGAGAGCC 1319
3879 CATAGACACAGACGTTCACTTCCAGGACAAATTAATGGCAGGCTCGGAAAAAATTAC 3938
1320 CATAGACACAGACGTTCACTTCCAGGACAAATTAATGGCAGGCTCGGAAAAAATTAC 1379
3939 TCCCGTAACCTGTATGTCATCAATGTCCTGCTGTACACATGATCCAGAAAA 3998
1380 TCCCGTAACCTGTATGTCATCAATGTCCTGCTGTACACATGATCCAGAAAA 1439
3999 CTTTGTACTACTGTATGTCATCAATGTCCTGCTGTACACATGATCCAGAAAA 4058
1440 CTTTGTACTACTGTATGTCATCAATGTCCTGCTGTACACATGATCCAGAAAA 1499
4059 CAAGTGAGCCGGAAGAAATGCTGGAAGCATTGTAAGGGGGTGTGTGCCCCACCAT 4118
1500 CAAGTGAGCCGGAAGAAATGCTGGAAGGTTGTGGGGGGTGTGTGCCCCACCAT 1559
4119 TCGTGCCCTGGGCAAGTCAAGTCCAGAAAGGTTGGAAGACCTTCTGTGAGGACAC 4178
1560 TCGTGCCCTGGGCAAGTCAAGTCCAGAAAGGTTGGAAGACCTTCTGTGAGGACAC 1619
4179 TTAATATCCCAAGAGTGGGTGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4238
1620 TTAATATCCCAAGAGTGGGTGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1679
4239 TCAATCCACAGAGTCAATCCGCGGTGCAACCAAGAGTCAATCCGCGGTGCAAC 4298
1680 TCAATCCACAGAGTCAATCCGCGGTGCAACCAAGAGTCAATCCGCGGTGCAAC 1739
4299 GAAATAGAGTCTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4358
1740 GAAATAGAGTCTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1799
4359 CTAAAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4418
1800 CTAAAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1859
4419 TACATGCTACCGCAGAGTGAATGTTACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 4478
1860 TACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1919
4479 TTTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4538
1920 TTTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1979
4539 CACATATAGT-----ATTTTGTGCTGAGCATTTGGCC 4571
1980 CCGATGATGTTTGAATGGAAGATCCAAATCCAGCACTGCTGCTGCTGCTGCTGCTGCT 2039
4572 AGAGGCACTGACAGTTGCCAGGGTGAAGTGAAGGGGCTGCTGCTGCTGCTGCTGCTGCT 4631
2040 GAGAGCACTGACAGTTGCCAGGGTGAAGTGAAGGGGCTGCTGCTGCTGCTGCTGCTGCT 2099
4632 AAATATCTTTTCAAGAGTCACTTCTGGGGTGTGGCTGTGCTGCTGCTGCTGCTGCTGCT 4691
2100 AAATATCTTTTCAAGAGTCACTTCTGGGGTGTGGCTGTGCTGCTGCTGCTGCTGCTGCT 2159
4692 GGTGTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4751
2160 GGTGTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2219
4752 TAATTTGAGCGGAGACAGAGTGAACA 4778
2220 TAATTTGAGCGGAGACAGAGTGAACA 2246

RESULT 6
US-08-643-219-12
: Sequence 12, Application US/08643219
: Patent No. 5801146
: GENERAL INFORMATION:
: APPLICANT: Davidson, Donald J.
: TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES
: TITLE OF INVENTION: AND METHODS FOR INHIBITING ANGIOGENESIS
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Abbott Laboratories
: STREET: 100 Abbott Park Road
: CITY: Abbott Park
: STATE: IL
: COUNTRY: USA
: ZIP: 60064
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/643,219
: FILING DATE: 06-MAY-1996
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Casuto, Dianne
: REGISTRATION NUMBER: 40,943
: REFERENCE/DOCKET NUMBER: 5940.US.P1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 847-938-3137
: TELEFAX: 847-938-2623
: TELEX:
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2497 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-643-219-12
Query Match 22.1%; Score 1104.8; DB 1; Length 2497;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 242; Indels 63; Gaps 2;
3159 TATCGAGGCATATCTCTCCAGCACTGTCTACAGGAAGACCTGTCAATCTGTGTCATCTATG 3218
896 TATCGCGGGAATGTGGCTGTTACCGGTGCCGGCACACACTGTGACACTGGAGTCACAG 955
3219 ATACCACTGGCATCAAGAGACCCAGAAAACTATCCCAATGCTGGCTGACCGAATAC 3278
956 ACCCTTCACACACATTAACAGGACACAGAAAACTTCCCTGCAAAAATTTGGATGAATAC 1015
3279 TACTGCAGGAATCCAGATTTTGGGAAACAACCTGTTGTACACAAACGATCCGTGTGTG 3338
1016 TACTGCCCAATCCGTACAGGAAAAAGGCCCCCATGTGTCATTAACCAACAGCAAGTGTG 1075
3339 AGTGGAGATCTGCAATCTGACACATGCTCAGAAACGATACAGTGTCTAGAGACT 3398
1076 CGGTGGAGTACTGTAAATATACCGTCTG----- 1104
3399 CCCACTGTGTTCCAGTTCCAAGCATGAGAGGCTATTCGTAAGACCAACCACTGAGCAA 3458
1105 -----TGACTCTCCCAATGATCCAGGACAAATTTGCTCCACAGCACACTGAGACTA 1159
3459 ACCCTGTGTGTCGGCAGTGTACCATGTATGAGGACAGAGTGTGAGGACATTTCTCC 3518
1160 ACCCTGTGTGTCGGCAGTGTACCATGTATGAGGACAGAGTGTGAGGACATTTCTCC 1219

3519 ACCACTGTACAGAGAGACATGTCAATCTTGGTCAATCATGACACACACCGGCATCAG 3578
1220 ACCACACACAG 1279
3579 AGAGACCCAG 3638
1280 AAGACCCAG 1339
3639 GCGCATACAG 3698
1340 GCGCATACAG 1399
3699 CTGACGAG 3758
1400 CTGACGAG 1459
3759 CCAAGCCTACAG 3818
1460 CCAAGCCTACAG 1519
3819 GCGACAG 3878
1520 GCGACAG 1579
3879 CATGACAG 3938
1580 CATGACAG 1639
3939 TGCCGTACAG 3998
1640 TGCCGTACAG 1699
3999 CTTTGTACAG 4058
1700 CTTTGTACAG 1759
4059 CAACTGAG 4118
1760 CAACTGAG 1819
4119 TCCGAG 4178
1820 TCCGAG 1879
4179 TTAATATCCAG 4238
1880 TTAATATCCAG 1939
4239 TCAATATCCAG 4298
1940 TCAATATCCAG 1999
4299 GAAATAG 4358
2000 GAAATAG 2059
4359 CTAAGCAG 4418
2060 CTAAGCAG 2119
4419 TACATGAG 4478
2120 TATGAG 2179
4479 TTTGAG 4538
2180 TTTGAG 2239
4539 CACTATAG 4571
2240 CCGATATAG 2299

4572 AGAGCAG 4631
2300 GAGAGCAG 2359
4632 AAATACATTTTACAG 4691
2360 AAATACATTTTACAG 2419
4692 GGTGTATGAG 4751
2420 GGTGTATGAG 2479
4752 TAATGAG 4769
2480 TAATGAG 2497

RESULT 7

US-09-131-995-12
Sequence 12 Application US/09131995
Patent No. 5972896

GENERAL INFORMATION:
APPLICANT: Davidson, Donald J.
TITLE OF INVENTION: NOVEL ANTIMIGRAINE PEPTIDES
TITLE OF INVENTION: AND METHODS FOR INHIBITING ANGIOGENESIS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/131,995
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/832,087
FILING DATE: 03-APR-1997
APPLICATION NUMBER: 08/643,219
FILING DATE: 06-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Casullo, Diane
REGISTRATION NUMBER: 40,943
REFERENCE/DOCKET NUMBER: 5940.US.P1
TELEPHONE: 847-938-3137
TELEFAX: 847-938-2623
TELEX:

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2497 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-131-995-12

Query Match 22.1%; Score 1104.8; DB 2; Length 2497;

Best Local Similarity 81.4%; Pred. No. 0; Mismatches 242; Indels 63; Gaps 2;

3159 TATGAG 3218
896 TATGAG 955
3219 ATACACAG 3278

Db 956 ACCCTTACACATTAACAGAGACCAAGAAACTTCCCTCGCAAAAATTGGATGAAAAAC 1015
Qy 3279 TACTGAGAAATCCAGATTTCGGGAAAACAACCTGGTGTACAAACCCATCCGTGTG 3338
Db 1016 TACTGCGCAATCCTGACGAAAAAGGGCCCATGGGCATACAAACCAAGCCAAAGTG 1075
Qy 3339 AGTGGAGTACTGCAATCTGACACAAATGCTCAGAAAACAGAAATCAGGTGTCTAGACT 3398
Db 1076 CGGTGGGAGTACTGTAAGATACCGTCTG----- 1104
Qy 3399 CCCACTGTGTCCAGTTCACAGATGAGGCTCATTTGAGAGCAGACCACTGAGCAA 3458
Db 1105 -----TACTTCCTCCCAATATCCAGGAACAATTGGCTCCAGACACCACTGAGCTA 1159
Qy 3459 ACCCTGTGTGTCGAGAGTCTACCATGTATATGGCCAGATTATGAGGACATTTCTCC 3518
Db 1160 ACCCTGTGTGTCGAGAGTCTACCATGTATGAGACAGCTACCGAGGACATCTCC 1219
Qy 3519 ACCACTGTACAGGAGAGACATGTCATTTGGTCATCCATGACACCAACCGGCATACAG 3578
Db 1220 ACCACACACACAGAGAAAGATGTCACTTTGGTCATCTATGACACCAACCGGACACAG 1279
Qy 3579 AGGACCCCAAGAAACTAACCAATGATGGCTGACAACTGAACTGACAGAAATCCAGAT 3638
Db 1280 AAGACCCCAAGAACTAACCAATGCTGGCTGACAACTGAACTGACAGAAATCCAGAT 1339
Qy 3639 GCCCATACAGAGCCCTGTGTGTATACATGACCCCAAGTCATGAGTGGAGTACTGCAAC 3698
Db 1340 GCCCATTAAGGCCCTGTGTGTATACACAGACCCCAAGTCATGAGTGGAGTACTGCAAC 1399
Qy 3699 CTGACGAGATGCTCAGACAGAGAGGAGTGTGCTGCTCCGATCTCATCCAGT 3758
Db 1400 CTGAAAAAATCTCAGAGAGAGAGAGAGTGTGTACACACTCCGCTGTGTCTGCT 1459
Qy 3759 CCAAGCTAGGGCTCTCTTGAACAAGACTGATGTTGGGAATGGGAAAGATACCGG 3818
Db 1460 CCAAGTGTAGAGACTCTCTCCGAAAGAGACTGATGTTGGGAATGGGAAAGATACCGA 1519
Qy 3819 GGCAGAGAGGACACACTGTACTGGGAGCCATGCCAGAAATGGGCTGCCAGGAGCC 3878
Db 1520 GGCAGAGAGGAGACACTGTACTGGGAGCCATGCCAGAAATGGGCTGCCAGGAGCC 1579
Qy 3879 CATAGACACAGACGTTTCATTCAGGAGCAAAATAAATGGGAGTCTGAAAAAAATTAC 3938
Db 1580 CATAGACACAGACATTTTCATTCAGGAGCAAAATCCAGGGGAGTCTGAAAAAAATTAC 1639
Qy 3939 TCCCGTAACTCTGATGGTACATCAATGTGCTGCTGCTACACATGAATCCAGAAA 3998
Db 1640 TCCCGTAACTCTGATGGTACATCAATGTGCTGCTGCTACACAAATCCAGAAA 1699
Qy 3999 CTTTGTGACTACTGATATCCCTCTCGTGACATCCCTCATTTGATTTGGGAAAGCT 4058
Db 1700 CTTTATCGACTACTGATATCCCTCTCATAGTGTGCGCCCTTCATTTGATTTGGGAAAGCT 1759
Qy 4059 CAAAGTGAAGCCGAAAGAAATGCTGGAAGCAATTTAGGGGGGTGTGGCCCAACCAT 4118
Db 1760 CAAAGTGAAGCCGAAAGAAATGCTGGAAGGTTTGAAGGGGTGTGGCCCAACCAT 1819
Qy 4119 TCCGAGCCCTGGCAAGTACGTCAGAAACAAGTTTGGAAAGCACTCTGTGAGAGGAC 4178
Db 1820 TCCGAGCCCTGGCAAGTACGTCAGAAACAAGTTTGAATGCACTTCTGTGAGGAGC 1879
Qy 4179 TTAATATCCCAAGATGGGTCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4238
Db 1880 TTAATATCCCAAGATGGGTCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1339
Qy 4239 TCATCTTCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4298
Db 1940 TCATCTTCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1999
Qy 4299 GAAATGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4358

Db 2000 GAAATGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2059
Qy 4359 CTAAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4418
Db 2060 CTAAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2119
Qy 4419 TACATGTCACCGCCAGACAGTCAATGTTACTACATGCTGCTGCTGCTGCTGCTGCTGCT 4478
Db 2120 TATGTCGCTGACCGAGCCGAATGTTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCT 2179
Qy 4479 TTTGGAGTGGCTCTTCAAGAGCCAGCTCCTGTTTATGAAATGAAATGCTGCAAT 4538
Db 2180 TTTGGAGTGGCTCTTCAAGAGCCAGCTCCTGTTTATGAAATGAAATGCTGCAAT 2239
Qy 4539 CACTTAAAT-----ATAATTGCTGAGCATTTGGCC 4571
Db 2240 CCGTATGAGTTTCAATGAAAGTCCAAATCCACCAACTCTGTGTGGCAATTTGGCC 2299
Qy 4572 AGAGCACTGACAGTGTGACAGGAGTACAGTGAAGGCTGCTGTTGCTTCAAGAGAC 4631
Db 2300 GAGGCACTGACAGTGTGACAGGAGTACAGTGAAGGCTGCTGTTGCTTCAAGAGAC 2359
Qy 4632 AATATCAATTTTCAAGAGTCACTTTGGGCTTGGCTGTGCAAGCCCAATTAAGCT 4691
Db 2360 AATATCAATTTTCAAGAGTCACTTTGGGCTTGGCTGTGCAAGCCCAATTAAGCT 2419
Qy 4692 GGTGTCTATGCTCGTGTTCAGAGTTTGTACTGATTTGAGGAGATGAGAAATAT 4751
Db 2420 GGTGTCTATGCTCGTGTTCAGAGTTTGTACTGATTTGAGGAGATGAGAAATAT 2479
Qy 4752 TAATTGACGGGAGACAG 4769
Db 2480 TAATTGACGGGAGACAG 2497

RESULT 8
US-08-832-087B-12
; Sequence 12, Application US/08832087B
; Patent No. 5981484
; GENERAL INFORMATION:
; APPLICANT: Davidson, Donald J.
; TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES
; TITLE OF INVENTION: AND METHODS FOR INHIBITING ANGIOGENESIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/832,087B
; FILING DATE: 03-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/643,219
; FILING DATE: 06-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Casuto, Dianne
; REGISTRATION NUMBER: 40,943
; REFERENCE/DOCKET NUMBER: 5940, US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-938-3137
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:

LENGTH: 2497 base pairs
Type: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-832-087B-12

Query Match 22.1%; Score 1104.8; DB 2; Length 2497;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 242; Indels 63; Gaps 2;

QY 3159 TATGAGGACATATCTCCACCACTGTACAGAGACCTGTCAATCTGTGTATATG 3218
DB 896 TATGCGGGATGTGCTGTATACCGTGTCCGACACACCTGTGAGCAGTGCACG 955
QY 3219 ATACCACTGCGATCAGAGACCCAGAAATACCAATCTGGCTGACGAGAAC 3278
DB 956 ACCCTCAGACACATTAACAGSACACGAAATACTCCCTGCAAAATTTGGATGAAC 1015
QY 3279 TACTGACAGAAATCCGATTCGTGGAAACAACCTGTGTACAAACCGATCCGTGTG 3338
DB 1016 TACTGCGCATCTCTGACGAAAGGGCCCATGTGCTACAAACAGCCACAGTG 1075
QY 3339 AGGTGGGATCTGCAATCTGACACATGCTCAGAAACAGAAATCAGGTGTCTAGAGCT 3398
DB 1076 CGGTGGGATCTGTAAATACCTCTG----- 1104
QY 3399 CCGACTGTTGTTCCAGTTCAGACGAGGCTCATCTGAGACGACCACTGACAA 3458
DB 1105 -----TGACTCTCCCGACGATCCACGGAACAAATGGCTCCACACACCTGAGCTA 1159
QY 3459 ACCCTGTGTCGCGAGTGTACATGTAATGGCCAGAGTATGAGGACATCTCC 3518
DB 1160 ACCCTGTGTCGAGATGTACATGTGTATGAGAGGCTACAGGACATCTCC 1219
QY 3519 ACCACTGTACAGAGAGACATCTGTGTATCTGACACACACCGGCTACG 3578
DB 1220 ACCACACACAGAGAAAGATGTGATCTGTGTATGACACACACCGGACAG 1279
QY 3579 AGGACCCAGAAATACCAATATGATGCTGACATGAACTACTGAGAAATCCAGAT 3638
DB 1280 AAGACCCAGAAATACCAATATGCTGCTGACATGAACTACTGAGAAATCCAGAT 1339
QY 3639 GCCCATACAGGCTCTGTGTATACATGACCCGACATCAGGTGAGATCTGCAAC 3698
DB 1340 GCCCATAAAGCCCTGTGTATACACAGACCCGACATGAGGTGAGATCTGCAAC 1399
QY 3699 CTGACGCGATCTCAGACACAGAGGAGTGTGCTGCTGCACTGATCAGAT 3758
DB 1400 CTGAAAAATATCTCAGAGAGAGAGGTGTGTACACTCGCTGTGTCTGCTT 1459
QY 3759 CCAGGCTAGGGCTCTCTTGAACAAGACTGTATGTTGGAAATGGAAAGATACCGG 3818
DB 1460 CCAGATGTAGAGACTCTCTCCGAAAGAGACTGTATGTTGGAAATGGAAAGATACCGA 1519
QY 3819 GGCAGAGAGGACACACCTGTTACTGGAGCGCATGCCAGGAATGGTCCGAGAGCC 3878
DB 1520 GGCAGAGAGGAGACACTGTTACTGGAGCGCATGCCAGAGATGGTCCGAGAGCC 1579
QY 3879 CATAGACACAGACGTTCACTTCAGGACAAATTAATGGCAGGTCTGAAAAAATTAAC 3938
DB 1580 CATAGACACAGACATTTTCACTTCAGAGACAAATTAATGGCAGGTCTGAAAAAATTAAC 1639
QY 3939 TGGCGTACCGCTGATGGGAGATCAATGTGCCGTGCTACCAATGATCCAAAGAAA 3998
DB 1640 TGGCGTAAACCTGATGGGAGATGAGGTGCCGTGCTACCAAGAAAATCCAAAGAAA 1699
QY 3999 CTTTGTGACTGTGATATCCCTCTCTGTGACATCTCTCATTTGATTTGGAGACCT 4058
DB 1700 CTTTACGACTGTGATATCCCTCTCTGTGAGGCGCCCTTCATTTGATTTGGAGACCT 1759
QY 4059 CAAGTGGAGCCGAGAAATGTCTGTGAGGACATTTGTAGGGGGGTGTGTGCGCCACCAT 4118
DB 1760 CAAGTGGAGCCGAGAAATGTCTGTGAGGAGGTTGTAGGGGGGTGTGTGCGCCACCAT 1819
QY 4119 TCTGGGCTCGGCGAGCTACATCTCAGACAGAGGTTTGGAAAGACTTCTGTGAGGACCC 4178
DB 1820 TCTGGGCTCGGCGAGCTACATCTCAGACAGAGGTTTGGAAAGACTTCTGTGAGGACCC 1879
QY 4179 TTAATATCCCGAGAGTGGTGTGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4238
DB 1880 TTAATATCCCGAGAGTGGTGTGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1939
QY 4239 TCACTCTACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4298
DB 1940 TCACTCTACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1999
QY 4299 GAAATAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4358
DB 2000 GAAATAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2059
QY 4359 CTAAGCAGGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4418
DB 2060 CTAAGCAGGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2119
QY 4419 TACATGCTACCGCGCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4478
DB 2120 TATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2179
QY 4479 TTTGGGAGTGGCTCTTCTCAAGAGAGCCAGCTCTTGTATGAGAAATGAGTGTGCT 4538
DB 2180 TTTGGGAGTGGCTCTTCTCAAGAGAGCCAGCTCTTGTATGAGAAATGAGTGTGCT 2239
QY 4539 CACTATAAGT-----AATTTGTGCTGAGATTTGGCC 4571
DB 2240 CGCTATGAGTTTGTGAATGGAAGATCCAAATCCAGCAATCTGTGCTGCTGCTGCTGCTGCTGCT 2299
QY 4572 AGAGGCACTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4631
DB 2300 AGAGGCACTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2359
QY 4632 AAATACATTTTACAGAGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4691
DB 2360 AAATACATTTTACAGAGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2419
QY 4692 GGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4751
DB 2420 GGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2479
QY 4752 TAATTGACGGAGAGACAG 4769
DB 2480 TAATTGACGGAGAGACAG 2497

RESULT 9

US-08-851-350-12

Sequence 12, Application us/08851350

Patent No. 6057122

GENERAL INFORMATION:

Applicant: Abbott Laboratories

TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES,

TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING SAME AND METHODS

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESS: Abbott Laboratories

STREET: 100 Abbott Park Road

CITY: Abbott Park

STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES
TITLE OF INVENTION: AND METHODS FOR INHIBITING ANGIOGENESIS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/132,154
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/832,087
FILING DATE: 03-APR-1997
APPLICATION NUMBER: 08/643,219
FILING DATE: 06-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Casuto, Dianne
REGISTRATION NUMBER: 40,943
REFERENCE/DOCKET NUMBER: 5940.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-938-3137
TELEFAX: 847-938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2497 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-132-154-12
Query Match 22.1%; Score 1104.8; DB 4; Length 2497;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 242; Indels 63; Gaps 2;
QY 3159 TATGAGGCAATATCTCCACACTGTACAGAGAGACCTGTCAATCTTGGTCATCTATG 3218
DB 896 TATGCGGGAATGAGCTGTACCGTGTCGGGACACACTGTACAGACTGGAGTGCACAG 955
QY 3219 ATACCACTGCGATCGAGAGACCCCAAACTACCCAAATGCTGGCTGCAGCAGAAC 3278
DB 956 ACCCTTACACACATTAACAGAGACCAAGAAATCTCCGCAAAATTTGGATGAAAC 1015
QY 3279 TACTGAGAGATCCAGATTTGTGGAAACAACCTGGTGTACACAAACCATCCGATGTG 3338
DB 1016 TATGCGGCAATCTGTAGCGAAGAAAGGCCCATGCTGCCATACCAACCAAGCAATG 1075
QY 3339 AGTGGAGATGACCAATCTGCACAACTCTGAGAAACAGAAATCAAGTGTCTTAGACT 3398
DB 1076 CGGTGGAGATCTGTAGATACCGTCTG----- 1104
QY 3399 CCCACTGTGTTCAGATTCACAGATGAGGCTCATTTGAAAGCAGACCAACTAGACAA 3458
DB 1105 -----TCACTCTCCACAGTATCCAGGAACAATTGGCTCCACAGCACCACCTTAGCTA 1159
QY 3459 ACCCTGTGTGCGGAGTGTACATGTAATGGCCAGATTTATGAGGACATTTCTCC 3518
DB 1160 ACCCTGTGTGCGGAGTGTACATGTAATGGGAGTGTAGAGAGTACCAGGACATCTCC 1219
QY 3519 ACCACTGTCAAGAGAGATGTCAATTTGGTCAATCATGACACACACCGGAGTACG 3578
DB 1220 ACCACACACAGAGAAAGTGTAGTCTGTGTCATCTATGACACACACCGGAGCAG 1279

QY 3579 AGACCCCAAGAACTACCCAAATGATGGCTGACAAATGAATCTACGAGGAATCCAGAT 3638
DB 1280 AAGACCCCAAGAACTACCCAAATGATGGCTGACAAATGAATCTACGAGGAATCCAGAT 1339
QY 3639 GCGGATACAGGCGCTGTGTGTACATGAGCCCAACATCAGTGGAGTATCTGCAC 3698
DB 1340 GCGGATACAGGCGCTGTGTGTACATGAGCCCAACATCAGTGGAGTATCTGCAC 1399
QY 3699 CTGACCGATGCTCAGACACAGAGAGTGTGTGCTCCCTCCGACTGTCCAGAGTT 3758
DB 1400 CTGAAAAATGCTCAGAGAACAGAGAGTGTGTAGCACTCCGCTGTGTCTCTCT 1459
QY 3759 CCAAGCTGAGGCTCTCTCTGACAAAGCTATGTTGGAAATGGGAAAGATACCAG 3818
DB 1460 CCAGATGAGACCTCTCCAGAGAAAGCTGTATGTTGGAAATGGGAAAGATACCGA 1519
QY 3819 GCGAAGACCAACACTGTACTGAGACCGCATCCAGAAATGGGCGCCCAAGACCC 3878
DB 1520 GCGAAGGCGGCGACACTGTACTGAGACCGCATCCAGAAATGGGCGCCCAAGACCC 1579
QY 3879 CATAGACACAGACGTTCAATCCAGGACAAATTAATGGCAGTCTGGAAAAAATTAC 3938
DB 1580 CATAGACACAGACGTTTCACTCCAGAGCAAAATCCAGGCGGCTGTGGAAAAAATTAC 1639
QY 3939 TGCCGTAACTGATGATGATCAATGATGATGATGATGATGATGATGATGATGATG 3998
DB 1640 TGCCGTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1699
QY 3999 CTTTGTACTGTGTATATCCCTCTGTGTGATCCTTTCAATTTGATTTGGAGACCT 4058
DB 1700 CTTTGTACTGTGTATATCCCTCTGTGTGATCCTTTCAATTTGATTTGGAGACCT 1759
QY 4059 CAAGTGAAGCGAAGAAATGCTCTGAAAGCATTTGAGGAGGAGTGTGAGCCACCCACAT 4118
DB 1760 CAAGTGAAGCGAAGAAATGCTCTGAAAGCATTTGAGGAGGAGTGTGAGCCACCCACAT 1819
QY 4119 TCCTGCGCTGTGCAATGCTCAGAACCAAGGTTTGGAAAGCACTTGTGAGGAGCC 4178
DB 1820 TCCTGCGCTGTGCAATGCTCAGAACCAAGGTTTGGAAAGCACTTGTGAGGAGCC 1879
QY 4179 TTAATATCCCAAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4238
DB 1880 TTAATATCCCAAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1939
QY 4239 TCATCTTACAGGATCATCTGCGGTGACACCAAGAAAGTGAATCTGATCTGATG 4298
DB 1940 TCATCTTACAGGATCATCTGCGGTGACACCAAGAAAGTGAATCTGATCTGATG 1999
QY 4299 GAAATGAAGTGTCTAGGCTGTCTTGGAGCCACACAGCAATATTCCTGCTGCTAAG 4358
DB 2000 GAAATGAAGTGTCTAGGCTGTCTTGGAGCCACACAGCAATATTCCTGCTGCTAAG 2059
QY 4359 CTAGAGAGGCTGTGCGGTCTATCTGACAAATGAATGCGAGTCTGCGATCCCGAC 4418
DB 2060 CTAGAGAGGCTGTGCGGTCTATCTGACAAATGAATGCGAGTCTGCGATCCCGAC 2119
QY 4419 TACATGTCACCGCAGGACTGATGTATCACTGTGCTGGGAGGAGAAACCAAGTAC 4478
DB 2120 TATGTGTGCTGACCGGACGCAATGTTTGTGCTGCTGGGAGGAGAAACCAAGTACT 2179
QY 4479 TTTGGAGCTGGCTTCTCAAGAAACCAAGTCTCTGTATTTGAGATGAAGTGTCAAT 4538
DB 2180 TTTGGAGCTGGCTTCTCAAGAAACCAAGTCTCTGTATTTGAGATGAAGTGTCAAT 2239
QY 4539 CACTATAAGT-----ATATTGTGCTGACATTTGGCC 4571
DB 2240 CCGTATGATTTCTGATGAGAAAGATCCAATCCACGAACTCTGTGCTGGCATTTGGCC 2299
QY 4572 AGAGGCACTGACAGTTGCCAGGAGTACAGTGAAGGCTCTGTGTTGCTTGAAGAGAC 4631
DB 2300 GAGGCACTGACAGTTGCCAGGAGTACAGTGAAGGCTCTGTGTTGCTTGAAGAGAC 2359
QY 4632 AATATATTTTCAAGAGATCACTTCTGTGGGCTCTGTGCTGTGACACCCCAATTAAGCCT 4691

```

|||||
Db 2360 AATACATTTTACAAAGAGTACTTCTGGGGTCTTGGCTGCGACGCCCAATTAAGCCT 2419
|||
QY 4692 GGTGCTATGCTCGTGTTCAGAGTTTGTACTTGATTTAGAGGAATGATGAATAAT 4751
|||||
Db 2420 GGTGCTATGCTCGTGTTCAGAGTTTGTACTTGATTTAGAGGAATGATGAATAAT 2479
|||
QY 4752 TAATTGAGCGAGACAG 4769
|||||
Db 2480 TAATTGAGCGAGACAG 2497
|||||

RESULT 11
5200340-5
; Patent No. 5200340
; APPLICANT: FOSTER, DONALD C.; MUIVHILL, EILEEN R.; O'HARA,
; PATRICK J.; PINDEL, KURT; YOSHITAKE, SHINJI
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
; ACTIVATORS
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/53,412
; FILING DATE: 22-MAY-1987
; SEQ ID NO:5:
; LENGTH: 1724

Query Match 14.0%; Score 697.6; DB 6; Length 1724;
Best Local Similarity 80.2%; Pred. No. 2,2e+208;
Matches 858; Conservative 0; Mismatches 179; Indels 33; Gaps 2;

QY 3786 GACTGATATTTGGATGGAGATGCGGGCAGAGAACCAACCTGTTACTGGG 3845
|||||
Db 655 GACTGCTACTTTGGATGGATGGCTACGCTTCCGTGCGACGACAGCCTCAGCGAGT 714
|||||
QY 3846 AGCCATGCCAGGAATGGCTGC-----CAAGAGCCCATAGACACAGCATTTCA 3899
|||||
Db 715 GCCTCCTGCTCCCGGTGGATTCATGATCTGATAGCGAAGTTTACACACAGAAC 774
|||||
QY 3900 CAGGAGCAAAATAATGGGAGGTCTGGAATAAATTAATGCGGTAAACCTGATGTC 3959
|||||
Db 775 CCGAGTCCCGAGGACTGGGCGCTGGCAACATTAATTCGCGGATCTGATGGGAT 834
|||||
QY 3960 ATCAATGTCCTGCTGTACACAAATGAAATCCAGAAAATTTTGTACTACTGTATATC 4019
|||||
Db 835 GCCAAGCCTGTGGTGCAGTGTGAAGAACCGAGCGTACGTGGGAGTACTGTATGTG 894
|||||
QY 4020 CCTCTGTGTCATCCTTTCATTTTATTTGGGAAGCTCAAGTGGAGCGGAAGATGT 4079
|||||
Db 895 CCTCCTGTGCGGTACCTTTCATTTGATTTGGGAAGCCCTCAAGTGGAGCGGAAGATGT 954
|||||
QY 4080 CCTGGAAGCATTTGAGGGGGTGTGGGCCACCCACATTTCTGGGCCCTGGCAAGTCA 4139
|||||
Db 955 CCTGGAAGGGTTGTAGGGGGGTGTGGGCCACCCACATTTCTGGGCCCTGGCAAGTCA 1014
|||||
QY 4140 CTCGAACAAGTTTGAAGAAGCTTCTGTGAGGCACTTATATCCCCAGAGTGGGTG 4199
|||||
Db 1015 CTGAACAAGTTTGAAGTCACTTCTGTGAGGCACTTATATCCCCAGAGTGGGTG 1074
|||||
QY 4200 CTGACTGCTGCTACTGCTTGAAGAAGTCTCAAGGCTTTCATCTCAAGAGTCACTG 4259
|||||
Db 1075 TTGACTGCTGCTGCTGCTTGAAGAAGTCCCAAGGCTTTCATCTCAAGAGTCACTG 1134
|||||
QY 4260 GGTGACACCAAGAAATGAACTCGAATCTCATGTTCAAGAAATGAAGTGTAGGCTG 4319
|||||
Db 1135 GGTGACACCAAGAAATGAACTCGAATCTCATGTTCAAGAAATGAAGTGTAGGCTG 1194
|||||
QY 4320 TTCTTGGAGCCACACAGAGATATTTGCTTGAAGCTTAAGCAGGCTCGCTCATTC 4379
|||||
Db 1195 TTCTTGGAGCCACACAGAAAGATATTTGCTTGAAGCTTAAGCAGGCTCGCTCATTC 1254
|||||
QY 4380 ACTGAAAGTATGCGACTTGTCTGCGATCCCGAGACTACATGCTACCGCGAGAGT 4439
|||||

```

```

Db 1255 ACTGCAAAAGTAAATCCAGCTTGTCTGCATCCCAATTAATGTGTGCTGACCGGACC 1314
|||
QY 4440 GAATGTTACATCACTGGCTGGGGAGAAACCAAGGTACTTTGGAGTGGCTCTTCAAG 4499
|||||
Db 1315 GAATGTTACATCACTGGCTGGGGAGAAACCAAGGTACTTTGGAGTGGCTCTTCAAG 1374
|||||
QY 4500 GAAAGCCAGCTCTGTTGATTAAGAAATGAAGTGTCAATCAATAAGT----- 4548
|||||
Db 1375 GAAGCCAGCTCCCTGTGATTAAGAAATGAAGTGTCAATCAATAAGTGTGAAATGA 1434
|||||
QY 4549 -----ATAATTTGCTGAGCATTTGGCCAGAGGCACTGACATGGCCAG 4592
|||||
Db 1435 AGAGTCCATCCACCAACTCTGTCTGGCTGGAGAGCAAAATACATTTTACAAGAGT 1494
|||||
QY 4593 GGTGACAGTGGAGGCGCTGTGCTGTGCTGAGAGCAAAATACATTTTACAAGAGT 4652
|||||
Db 1495 GGTGACAGTGGAGTCTCTGTGCTGTGCTGAGAGCAAAATACATTTTACAAGAGT 1554
|||||
QY 4653 ACTTCTGGGGCTTGTGCTGTGCAAGCCCAATTAAGCTGTGTATGCTGCTGTTCA 4712
|||||
Db 1555 ACTTCTGGGGCTTGTGCTGTGCAAGCCCAATTAAGCTGTGTATGCTGCTGTTCA 1614
|||||
QY 4713 AGTTTGTACTTGTGATTTAGAGGAATGATGAGAAATTAATTAATTTAGCGGAGACAGT 4772
|||||
Db 1615 AGTTTGTACTTGTGATTTAGAGGAATGATGAGAAATTAATTAATTTAGCGGAGACAGT 1674
|||||
QY 4773 GAAGATCAACCTACTTGAAGCTGAACGTTGGGTAAGGATTTAGCATGC 4822
|||||
Db 1675 GACGCACTGACTCACCTTAGAGCTGGAAACGAGGTAGGATTTAGCATGC 1724
|||||

RESULT 12
US-08-379-621-1
; Sequence 1, Application US/08379621
; Patent No. 5645833
; GENERAL INFORMATION:
; APPLICANT: DAWSON, Keith
; APPLICANT: GILBERT, Richard James
; TITLE OF INVENTION: INHIBITOR RESISTANT SERINE PROTEASES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HALE AND DORR
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379, 621
; FILING DATE: 03-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/01632
; FILING DATE: 03-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BAKER, Hollie L.
; REGISTRATION NUMBER: 31,321
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 942-8400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 690 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:

```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..690
; OTHER INFORMATION: /partial
; OTHER INFORMATION: /codon_start=1
; OTHER INFORMATION: /function="encodes plasmin protease domain"
; OTHER INFORMATION: /product="nucleotide with corresponding
; OTHER INFORMATION: protein"
; OTHER INFORMATION: /number=1
US-08-379-621-1

Query Match      11.2%; Score 557.8; DB 1; Length 690;
Best Local Similarity 90.0%; Pred. No. 8.6e-165;
Matches 620; Conservative 0; Mismatches 42; Indels 27; Gaps 1;

QY 4090 TTGTAGGGGGGTGTGTGGCCACCACCATTCCTGGCCCTGGCAAGTCAGTCTAGAACAA 4149
    |||||
DB 2 TTGTAGGGGGGTGTGTGGCCACCACCATTCCTGGCCCTGGCAAGTCAGTCTAGAACAA 61

QY 4150 GGTGGAAAGCACTTCTGTGAGGACCTTAATATCCAGAGTGGGTCTGACTGCTG 4209
    |||||
DB 62 GGTGGAAAGCACTTCTGTGAGGACCTTGAATATCCAGAGTGGGTCTGACTGCTG 121

QY 4210 CTCAGTCTTGAAGAACTCTCAAGGCTTCATCTACAGATCATCTGGTGCACACC 4269
    |||||
DB 122 CCCACTCTTGAGAAAGTCCCAAGGCTTCATCTACAGATCATCTGGTGCACACC 181

QY 4270 AAGAAGTAAGCTCGAATCTCATGTTTCAGAAATAGAGTGTAGGCTTTTGGAGC 4329
    |||||
DB 182 AAGAAGTAAGCTCGAAGCCAGTGTGAGAAATAGAGTGTAGGCTTTTGGAGC 241

QY 4330 CCACAGCAAGCAATATTCCTTCAAGGTAAGAGGCTGCGCATCATCTGACAAAG 4389
    |||||
DB 242 CCACAGCAAGCAATATTCCTTCAAGGTAAGAGTGTGCGCATCATCTGACAAAG 301

QY 4390 TAATGCCAGCTTCTGTCACATCCCAAGTACATGTTACCGCCAGAGTGAATGTTACA 4449
    |||||
DB 302 TAATGCCAGCTTCTGTCACATCCCAAGTATATGTTGCGGTGACCGAGCAATGTTACA 361

QY 4450 TCACGTGGCTGGGAGAAACCCAGGATACCTTTGGAGTGGCTTCTCAAGAAAGCCAGC 4509
    |||||
DB 362 TCACGTGGCTGGGAGAAACCCAGGATACCTTTGGAGTGGCTTCTCAAGAAAGCCAGC 421

QY 4510 TCCTTTATTTAGAAATGAGTGTGCAATCACTAATAGT----- 4548
    |||||
DB 422 TCCTTTATTTAGAAATGAGTGTGCAATCACTAATAGT----- 481

QY 4549 -----ATATTTGTGCTGAGCATTTGGCCAGAGGACACTGACAGTGGTGCAGTG 4602
    |||||
DB 482 CCACGCAACTCTGTGCTGGGCAATTTGGCCGAGGACACTGACAGTGGTGCAGTG 541

QY 4603 GAGGGCTCTGCTTGTCTTGTGAGAAAGCAAAATACATTTTACAGAGATCACTTCTGGG 4662
    |||||
DB 542 GAGGGCTCTGCTTGTCTTGTGAGAAAGCAAAATACATTTTACAGAGATCACTTCTGGG 601

QY 4663 GTCTGGCTGTGACGCCCCCAATAGCTGTGATGCTGCTGTTTCAAGTTTGTGA 4722
    |||||
DB 602 GTCTGGCTGTGACGCCCCCAATAGCTGTGATGCTGCTGTTTCAAGTTTGTGA 661

QY 4723 CTTCGATTGAGGGAATGATGAGAAATAT 4751
    |||||
DB 662 CTTCGATTGAGGGAATGATGAGAAATAT 690

RESULT 13
US-08-889-078-1
; Sequence 1, Application US/08889078
; Patent No. 5932213
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith Marilyn
; APPLICANT: Richard James Gilbert
; TITLE OF INVENTION: MODIFIED PLASMIN PRECURSORS WITH RESISTANCE TO
```

```

; TITLE OF INVENTION: INHIBITORS OF PLASMIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HALE AND DORR LLP
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: United States of America
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/889,078
; FILING DATE: 7 JULY 1997
; CLASSIFICATION: 536
; NAME: Baker, Hollie L.
; REGISTRATION NUMBER: 31,321
; REFERENCE/DOCKET NUMBER: 102378.187CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 526-6000
; TELEFAX: (617) 526-5000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 690 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..690
; OTHER INFORMATION: /partial
; OTHER INFORMATION: /codon_start=1
; OTHER INFORMATION: /function="encodes plasmin protease domain"
; OTHER INFORMATION: /product="nucleotide with corresponding
; OTHER INFORMATION: protein"
; OTHER INFORMATION: /number=1
US-08-889-078-1

Query Match      11.2%; Score 557.8; DB 2; Length 690;
Best Local Similarity 90.0%; Pred. No. 8.6e-165;
Matches 620; Conservative 0; Mismatches 42; Indels 27; Gaps 1;

QY 4090 TTGTAGGGGGGTGTGTGGCCACCACCATTCCTGGCCCTGGCAAGTCAGTCTAGAACAA 4149
    |||||
DB 2 TTGTAGGGGGGTGTGTGGCCACCACCATTCCTGGCCCTGGCAAGTCAGTCTAGAACAA 61

QY 4150 GGTGGAAAGCACTTCTGTGAGGACCTTAATATCCAGAGTGGGTCTGACTGCTG 4209
    |||||
DB 62 GGTGGAAAGCACTTCTGTGAGGACCTTGAATATCCAGAGTGGGTCTGACTGCTG 121

QY 4210 CTCAGTCTTGAAGAACTCTCAAGGCTTCATCTACAGATCATCTGGTGCACACC 4269
    |||||
DB 122 CCCACTCTTGAGAAAGTCCCAAGGCTTCATCTACAGATCATCTGGTGCACACC 181

QY 4270 AAGAAGTAAGCTCGAATCTCATGTTTCAGAAATAGAGTGTAGGCTTTTGGAGC 4329
    |||||
DB 182 AAGAAGTAAGCTCGAAGCCAGTGTGAGAAATAGAGTGTAGGCTTTTGGAGC 241

QY 4330 CCACAGCAAGCAATATTCCTTCAAGGTAAGAGGCTGCGCATCATCTGACAAAG 4389
    |||||
DB 242 CCACAGCAAGCAATATTCCTTCAAGGTAAGAGTGTGCGCATCATCTGACAAAG 301

QY 4390 TAATGCCAGCTTCTGTCACATCCCAAGTACATGTTACCGCCAGAGTGAATGTTACA 4449
    |||||
```

```
Dh 302 TAATCCAGCTTGTCTGCATCCCAATTAATGTGTGCTGACCGGAGCCGAATGTTTCA 361
Qy 4450 TCACTGGCTGGGAGAAAACCAAGTACCTTTGGAGTGGCTTCTCAAGAAAGCCGAC 4509
Dh 362 TCACGTGGCTGGGAGAAAACCAAGTACCTTTGGAGTGGCTTCTCAAGAAAGCCGAC 421
Qy 4510 TCCCTGTATAGATGAAGTGTGCAATCACTATTAAGT-----4548
Dh 422 TCCCTGTATAGATGAAGTGTGCAATCACTATTAAGT-----4548
Qy 4549 -----ATATTTGTGCTGAGCATTTGGCCAGAGCCACTGACAGTTGCCAGGGTGCACGTG 4602
Dh 482 CCACCGAATCTGTGTGGGCAATTTGGCCAGAGCCACTGACAGTTGCCAGGGTGCACGTG 541
Qy 4603 GAGGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4662
Dh 542 GAGGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 601
Qy 4663 GTCTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4722
Dh 602 GTCTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 661
Qy 4723 CTTGGATTGAGGGGAATGATGAAATTAAT 4751
Dh 662 CTTGGATTGAGGGGAATGATGAAATTAAT 690
```

RESULT 14

```
US-09-206-059-29
; Sequence 29, Application US/09206059
; Patent No. 6201104
; GENERAL INFORMATION:
; APPLICANT: Macdonald, Nicholas
; TITLE OF INVENTION: Angiogenesis-Inhibiting Protein Binding Peptides and
; FILE REFERENCE: 05213-0370
; CURRENT APPLICATION NUMBER: US/09/206,059
; CURRENT FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-206-059-29
```

Query Match 5.1%; Score 257.2; DB 4; Length 1134;

Best Local Similarity 66.8%; Pred. No. 2.4e-70;

Matches 401; Conservative 0; Mismatches 163; Indels 36; Gaps 1;

```
Qy 3159 TATCGAGGCAATATCCACACACATGTCACAGGAAGCACTGTCATCTGGTCATCTATG 3218
Dh 571 TATCGAGGCAATATCCACACACATGTCACACACCTGTCAGCATGAGTGCACAG 630
Qy 3219 ATACACACTGTCATGAGAGGACCCAGAAAATCACTACCAATATGCTGCTGACCGAAGAC 3278
Dh 631 ACCCTCAGACACATATAACAGAGAACCAAAAATTTCCCGCAAAAATTTGGATGAAAC 690
Qy 3279 TACTGCAGAGATCCAGATTTCTGGGAAAACCTGTGTTCACAAACGATCCGTTGTG 3338
Dh 691 TACTGCAGAGATCCAGATTTCTGGGAAAACCTGTGTTCACAAACGATCCGTTGTG 750
Qy 3339 AGGTGGAGATGACGATGACACAAATGCTCAGAAAAGAAACAGATGCTCAGAGACT 3398
Dh 751 CGGTGGAGATGACGATGACACAAATGCTCAGAAAAGAAACAGATGCTCAGAGACT 779
Qy 3399 CCCACTGTGTGTCAGTTCCAAACATGAGGGCTCATTTGAAAGCAGCACCACCACTGAGCAA 3458
Dh 780 -----TACTGCTCCCAAGATATCCACAGAAACATTTGGCTCCACACACACCTGAGCTA 834
Qy 3459 ACCCTGTGTGTCGCGAGTGTCTACATGTTAATGCGCAGAGTTATGAGGAGACATTTCTCC 3518
Dh 834 ACCCTGTGTGTCGCGAGTGTCTACATGTTAATGCGCAGAGTTATGAGGAGACATTTCTCC 3518
```

```
Dh 835 ACCCTGTGTGTCGCGAGTGTCTACATGTTAATGAGACAGACTACCGAGGACACATCTCC 894
Qy 3519 ACCACTGTACAGAGAGACATGTCATCTTGTGTCATCCATGACACACACCGGCAATCAG 3578
Dh 895 ACCACACACACAGAGAAAGAGTGTGACGTGTGTCATCTATGACACACACCGGCAATCAG 954
Qy 3579 AGACCCCGAGAAAATCTACCCAAATGATGGCTGTGACATGAACTACTGACAGAAATCCAGAT 3638
Dh 955 AAGACCCCGAGAAAATCTACCCAAATGATGGCTGTGACATGAACTACTGACAGAAATCCAGAT 1014
Qy 3639 GCCGATACAGGCCCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3698
Dh 1015 GCCGATACAGGCCCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1074
Qy 3699 CTGACGATGATCTGACACACAGAGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3758
Dh 1075 CTGAAAAATGCTGACGAAACAGAACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1134
```

RESULT 15

```
US-08-030-410-2
; Sequence 2, Application US/08030410
; Patent No. 6221359
; GENERAL INFORMATION:
; APPLICANT: Komiyama, Atsushi
; APPLICANT: Nakahata, Tatsutoshi
; APPLICANT: Kubo, Tetsuo
; APPLICANT: Tanaka, Ryuhel
; APPLICANT: Kawano, Genji
; APPLICANT: Sano, Emiko
; APPLICANT: Kojima, Katsuki
; TITLE OF INVENTION: HEMOPOIETIC STEM CELL MULTILIBR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/030,410
; FILING DATE: 19930521
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digilio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8898
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2172 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2169
; US-08-030-410-2
```

Query Match 1.8%; Score 90.8; DB 4; Length 2172;

Best Local Similarity 50.5%; Pred. No. 5.7e-18;

Matches 279; Conservative 0; Mismatches 237; Indels 36; Gaps 1;

QY 2130 GATTGCTACCATGGTGTGACAGAGTTATCGAGGCTCATTTCTTAACACTGTCAACAGCA 2189
|| ||| | ||| ||||| ||||| | ||| |||||
Db 613 GAATGCAATGACCTGCAATGGGAGAGTTATCGAGGTCTCATGGATCATACAGAAATCAGGC 672
QY 2190 AGGACATGTGACGCTGCTGCTATATGACACCACTGGCATGCATGAGAGCAACAGAAATAT 2249
||| ||||| ||| | ||||| ||||| ||| | |||
Db 673 AGGATTTGTGACGCGCTGGATTCATCAGACACACCGGCAAAATTTCTGCTGAAGA 732
QY 2250 TATCCAAATGGTGCCCTGACAGAGAACTACTGAGAAATCCAGATGCTGAGATTAGTCT 2309
||| | ||| | ||| | ||| | ||| | ||| | |||
Db 733 TATCCGACAAAGGCGCTTGAATGAATAATTTATGCGCAATCCGATGCGCCAGCGAGCCA 792
QY 2310 TGGTGTATATACATGGATCCCAATGTCAATGGGAGAGTACTGCAACTGACACAATGTCCA 2369
||| ||||| ||| | ||| | ||||| ||||| | ||| |
Db 793 TGGTGTATATCTTTGACCTTCACACCGGCTGGGAGTACTGTCAATTAATAACATGCGCT 852
QY 2370 GTGACAGAAATCAAGTGTCTTGGACGCTCCAGGCTGTCTTGAACAAGCACCAACGAG 2429
| | | | | | | | | | | | | | | | | | |
Db 853 GACAAATCTATGAAATGACACTGATGTCTTTGGAAACAACTGAA----- 897
QY 2430 CAAAGCCCCACAGTCCAGACCTGCTACCATGTGTATGACAGAGATTATCGAGGCTCATTC 2489
||| ||| ||| | ||| | ||| | ||| | ||| |
Db 898 -----TGCATCCAAAGTCAAGAGAGAGGCTTACAGGGGCACTGTC 936
QY 2490 TCCACCACTGTACAGAGAGGACATGTCACTTGGTCTCTATGACACCACTGGCAT 2549
||| | | ||||| ||||| ||||| ||| | ||||| |||||
Db 937 AATACCATTTGGAAATGSAATTCATGTCAAGCTTGGGATTTCTCAGTATCTCTCAGAGCAT 996
QY 2550 CAGAGAAACACAGAAATCTACCCAAATGTGGCTGACAGAACTACTGACAGAAATCCA 2609
| | | | | | | | | | | | | | | | | | |
Db 997 GACATGACTCTCTGAATAATTCAGTGCAAGACCTACGAGAAATTAATCTCCGAAATCCA 1056
QY 2610 GATGCTGAGATTCGCTTGGTGTATACCATGATCCCGTGCAGATGGAGTACTGC 2669
||| ||||| ||||| ||||| ||||| ||||| | ||| |
Db 1057 GATGGGTCTGATCAACCTGGTGTTTACCACTGATCCAAACATCCGAGTTGGCTACTGC 1116
QY 2670 AACCTGACGCAA 2681
| | | | |
Db 1117 TCCCAAAATCCA 1128
| | | | |

Search completed: March 6, 2003, 15:26:56
Job time : 273 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 06:40:43 ; Search time 758 Seconds
(without alignments)
4405.804 Million cell updates/sec

Title: US-09-923-515-3_COPY_8939_13938

Perfect score: 5000

Sequence: 1 ctcgcactgttaccgccggtt.....ctgcactatttgatttga 5000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 478924 seqs, 333959956 residues

Total number of hits satisfying chosen parameters: 957848

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications -NA:*
1: /cgn2_6/pdata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/pdata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/pdata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/pdata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/pdata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/pdata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/pdata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/pdata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/pdata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/pdata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/pdata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/pdata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/pdata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/pdata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3753.2	75.1	10422	9	US-09-870-759-125
2	1106.4	22.1	2497	10	US-09-946-893-1
3	654.6	13.1	1907	9	US-09-924-340-53
4	654.6	13.1	1907	9	US-09-992-600A-53
5	306.4	6.1	1236	10	US-09-880-107-2341
6	257.2	5.1	1134	10	US-09-873-676-29
7	235.6	4.7	384	10	US-09-960-352-10993
8	231.4	4.6	468	10	US-09-960-352-10851
9	229.2	4.6	372	10	US-09-960-352-7521
10	204.4	4.1	313	9	US-09-796-692-5593
11	197	3.9	394	10	US-09-960-352-11088
12	193.8	3.9	346	10	US-09-960-352-12743
13	145.2	2.9	422	10	US-09-833-381-252
14	139.6	2.8	234	10	US-09-960-352-14383
15	134.8	2.7	199	10	US-09-960-352-11543
16	134.8	2.7	431	10	US-09-864-761-13941
17	129.6	2.6	160	10	US-09-864-761-20992
18	122.4	2.4	408	10	US-09-960-352-1535
19	108.4	2.2	203	10	US-09-960-352-11298

20	104.2	2.1	168	10	US-09-864-761-30505	Sequence 30505, A
c	94	1.9	362	10	US-09-864-761-4239	Sequence 4239, Ap
21	93.4	1.9	242	10	US-09-960-352-12971	Sequence 12971, A
22	89.2	1.8	1344	9	US-10-081-309-1	Sequence 1, Appl1
23	89.2	1.8	5898	10	US-09-880-107-3708	Sequence 3708, Ap
24	89.2	1.6	645	12	US-10-036-869-22	Sequence 22, Appl
25	80.4	1.6	1284	12	US-10-036-869-24	Sequence 24, Appl
26	80.4	1.5	780	9	US-09-870-759-96	Sequence 96, Appl
27	77.2	1.5	402	10	US-09-960-352-15108	Sequence 15108, A
28	71.6	1.4	792	10	US-09-888-615-37	Sequence 37, Appl
29	68	1.4	614	10	US-09-879-792-33	Sequence 33, Appl
30	67.6	1.4	1222	9	US-10-125-459-3	Sequence 3, Appl1
31	66.6	1.3	1222	9	US-10-125-459-5	Sequence 5, Appl1
32	66.6	1.3	1222	9	US-10-067-761-5	Sequence 5, Appl1
33	66.6	1.3	1222	10	US-09-946-633-3	Sequence 3, Appl1
34	66.6	1.3	1230	10	US-09-879-792-35	Sequence 35, Appl
35	66	1.3	1614	10	US-09-888-615-45	Sequence 45, Appl
36	66	1.3	1748	10	US-09-879-792-11	Sequence 11, Appl
37	66	1.3	1079	9	US-10-125-459-5	Sequence 5, Appl1
38	64.6	1.3	1079	9	US-10-067-761-7	Sequence 7, Appl1
39	64.6	1.3	1079	10	US-09-804-156-7	Sequence 7, Appl1
40	64.6	1.3	1079	10	US-09-946-633-5	Sequence 5, Appl1
41	64.6	1.3	1079	10	US-10-125-459-4	Sequence 4, Appl1
42	64.6	1.3	1394	9	US-10-067-761-6	Sequence 6, Appl1
43	64.6	1.3	1394	10	US-09-804-156-6	Sequence 6, Appl1
44	64.6	1.3	1394	10	US-09-946-633-4	Sequence 4, Appl1
45	64.6	1.3	1394	10	US-09-946-633-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-870-759-125
; Sequence 125, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 125
; LENGTH: 10422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6667)..(10176)
; OTHER INFORMATION:
US-09-870-759-125

Query Match 75.1%; Score 3753.2; DB 9; Length 10422;
Best local similarity 99.8%; Pred. No. 0;
Matches 3758; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1235 TGGCTTGATCAGTACTGACGAGATCCAGATCCTGTGACGCCCTTATTGTATAC 1294
|||||
Db 6657 TGGCTTGATCAGTACTGACGAGATCCAGATCCTGTGACGCCCTTATTGTATAC 6716
|||||
QY 1295 GAGGGATCCAGTGTGAGTGGAGATCGCAACCTGCACACATGCTGAGACCGAGAG 1354
|||||
Db 6717 GAGGGATCCAGTGTGAGTGGAGATCGCAACCTGCACACATGCTGAGACCGAGAG 6776
|||||
QY 1355 GACTGCGTGGCGCCCTCCACTATTACCCCGATTCCAGAGCTTAGAGGCTCTTGACAA 1414
|||||
Db 6777 GACTGCGTGGCGCCCTCCACTATTACCCCGATTCCAGAGCTTAGAGGCTCTTGACAA 6836
|||||
QY 1415 AGACCACTGACGCAAGGCTGGGGTGCAGAGAGTGCTACCGAATGACAGACTTA 1474
|||||

Db	6837	AGCACACACTGAGCAAAAGGCCCTGGGGTGCAGAGTGCTACCAAGAAATGACAGAGTTA	6896
QY	1475	TCGAAGGACATACTCTTCACTTACTGTCAAGAGAAACCTCCCAAGCTTGTCATCTATGAC	1534
Db	6897	TCAGAGCACATACTTCAATTACTGTCAACAGAAAGAACCTGCCAAGCTTGGTATCTATGAC	6956
QY	1555	ACCACTACTGCCATAGTGTGGACCCCAACCATCTACCCAAATGCTGGCTGATCAAGACTA	1594
Db	6957	ACCACTCTCCATATGTGACCCCAACCATCTACCCAAATGCTGGCTGATCAAGACTA	7016
QY	1595	CTGCCGAATATCCAGATCTCTGGCAGCCCTTGTTGTATACAAAGATCCAGGTCTAG	1654
Db	7017	CTGCCGAATATCCAGATCTCTGGCAGCCCTTGTTGTATACAAAGATCCAGGTCTAG	7076
QY	1655	GTGGGAGTACTGCACACTGACAGATGCTCAGATGCAATGAGATGGACTGCTTGCTCTCC	1714
Db	7077	GTGGGAGTACTGCACACTGACAGATGCTCAGATGCAATGAGATGGACTGCTTGCTCTCC	7136
QY	1715	GAATGTTATCTCGCTCCAGCCTTAAGAGCTTTTGTGAACAGACAGTACGAGAGAAC	1774
Db	7137	GAATGTTATCTCGCTCCAGCCTTAAGAGCTTTTGTGAACAGACAGTACGAGAGAAC	7196
QY	1775	CCCCGGGGTACAGAGCTGCTACTACCATTAATGACAGAGTTACCGAGGCATACTCTCAC	1834
Db	7197	CCCCGGGGTACAGAGCTGCTACTACCATTAATGACAGAGTTACCGAGGCATACTCTCAC	7256
QY	1835	CACGTGACAGAGAAACTTGCCCAACCTTGTCATCTATGACACACACAGCATATGCG	1894
Db	7257	CACGTGACAGAGAAACTTGCCCAACCTTGTCATCTATGACACACACAGCATATGCG	7316
QY	1895	GACCCCAAGAAACTACCCCAATGCTGGCTGACACAGAACTACTGACAGAAATCCAGATGC	1954
Db	7317	GACCCCAAGAAACTACCCCAATGCTGGCTGACACAGAACTACTGACAGAAATCCAGATGC	7376
QY	1955	TGAGATTTCGGCTCTGGGTGTACACCATGGATCCCAAGTGCAGTGGGGTACTGCACACT	2014
Db	7377	TGAGATTTCGGCTCTGGGTGTACACCATGGATCCCAAGTGCAGTGGGGTACTGCACACT	7436
QY	2015	GACACAAATGCTGTGACAGAAATCAAGTGTCTTGCACTCACAGGAGTCCCAAGATCC	2074
Db	7437	GACACAAATGCTGTGACAGAAATCAAGTGTCTTGCACTCACAGGAGTCCCAAGATCC	7496
QY	2075	AAGACAGAGAGGCTTCTCTGSAABAACACCAAGACAGCAAAACCCCGGGGTCCAGATTCG	2134
Db	7497	AAGACAGAGAGGCTTCTCTGSAABAACACCAAGACAGCAAAACCCCGGGGTCCAGATTCG	7556
QY	2135	CTACCATGATGATGACAGAGATTTGAGGCTCATTTCTTACACATGTCACAGAAAGAC	2194
Db	7557	CTACCATGATGATGACAGAGATTTGAGGCTCATTTCTTACACATGTCACAGAAAGAC	7616
QY	2195	ATGCGACTCTTGCTCTCTATGACACCAACATGSGCATCAGAGAGCAACAGAAATATTATCC	2254
Db	7617	ATGCGACTCTTGCTCTCTATGACACCAACATGSGCATCAGAGAGCAACAGAAATATTATCC	7676
QY	2255	AAATGATGAGCTTACACAGAACTACTGCGAGATCCAGATGCTGAGATTATGCTTGGTG	2314
Db	7677	AAATGATGAGCTTACACAGAACTACTGCGAGATCCAGATGCTGAGATTATGCTTGGTG	7736
QY	2315	TTATATCATATGGATCCCAATATGTCAGATGGAGTACTGCAACTGACACAAATGTCAGTATC	2374
Db	7737	TTATATCATATGGATCCCAATATGTCAGATGGAGTACTGCAACTGACACAAATGTCAGTATC	7796
QY	2375	AGAAATCAAGTGTCTTGCGACGCGCAACGGCTGTCTTCTAACAAGCAACAGGAGCAAG	2434
Db	7797	AGAAATCAAGTGTCTTGCGACGCGCAACGGCTGTCTTCTAACAAGCAACAGGAGCAAG	7856
QY	2435	CCCCACAGTCCAGAGCTGCTACCATGGTATGAGACAGATTATCCAGGCTCATTTCTCCAC	2494
Db	7857	CCCCACAGTCCAGAGCTGCTACCATGGTATGAGACAGATTATCCAGGCTCATTTCTCCAC	7916
QY	2495	CACGTGTTACAGGAAGACATGTCAGTCTTGGTCTCTATGACACCACACTGGCATGAG	2554
Db	7917	CACGTGTTACAGGAAGACATGTCAGTCTTGGTCTCTATGACACCACACTGGCATGAG	7976

QY	2555	AACCAACAATACCTACCAAAATGTTGGCTTCAGCCAGGAACCTACTGCAAGAACTCCAGATGC	2614
Db	7977	AACCAACAATAATACCAAAATGTTGGCTTCAGCCAGGAACCTACTGCAAGAACTCCAGATGC	8038
QY	2615	TGAGATTGGCCCTTGTTGTTATACATGATGATCCAGTGTACATGGAGTAATCTGCACACT	2674
Db	8037	TGAGATTGGCCCTTGTTGTTATACATGATGATCCAGTGTACATGGAGTAATCTGCACACT	8098
QY	2675	GACGCAATGTCCAGTGAATGGAATCAGTCTCTCTCAGCACTCCAGTGGTGTCCAGTTC	2734
Db	8097	GACGCAATGTCCAGTGAATGGAATCAGTCTCTCTCAGCACTCCAGTGGTGTCCAGTTC	8156
QY	2735	AAGCAGAGCTTCCTCTTGAGAGCAGCCACATGTGAAAAGACACTGGGGTGTCCAGACTG	2794
Db	8157	AAGCAGAGCTTCCTCTTGAGAGCAGCCACATGTGAAAAGACACTGGGGTGTCCAGACTG	8218
QY	2795	CTACCGAGTGATGAGACAGATTAATGAGGACACACTCTCCACACATACACAGGAAGAC	2854
Db	8217	CTACCGAGTGATGAGACAGATTAATGAGGACACACTCTCCACACATACACAGGAAGAC	8278
QY	2855	ATGTCAGTCTTGTCGTCTATGACACACCAATTTGGATCGAGAGATCCACTTATCTATCC	2914
Db	8277	ATGTCAGTCTTGTCGTCTATGACACACCAATTTGGATCGAGAGATCCACTTATCTATCC	8336
QY	2915	AAATCGTGGCCTTGACCAAGAACTACCTCAGAGATCCAGATGCGTAGATTCGCCCTTGGTG	2974
Db	8337	AAATCGTGGCCTTGACCAAGAACTACCTCAGAGATTCAGATGCGTAGATTCGCCCTTGGTG	8396
QY	2975	TTACACCAATGGATCCCAAGTGTCAAGSTGGAGTACTGCAACTGTACACGATGTCCAGTAC	3034
Db	8397	TTACACCAATGGATCCCAAGTGTCAAGSTGGAGTACTGCAACTGTACACGATGTCCAGTAC	8456
QY	3035	AGAAATCGAGTTCCTTCACCACTCCACAGTGGCCCGGTTCCAAAGACACAGGGCTCTTC	3094
Db	8457	AGAAATCGAGTTCCTTCACCACTCCACAGTGGCCCGGTTCCAAAGACACAGGGCTCTTC	8516
QY	3095	TGAACAAGCACACCTGAGAAAAAGCCCTGTGTCCAGATTCCTACCATGTGTATGACG	3154
Db	8517	TGAACAAGCACACCTGAGAAAAAGCCCTGTGTCCAGATTCCTACCATGTGTATGACG	8576
QY	3155	GAGTTATGAGGCAATTCCTCCACCACTGTCAAGAGAAAGACTGTGCATCTTGCTATC	3214
Db	8577	GAGTTATGAGGCAATTCCTCCACCACTGTCAAGAGAAAGACTGTGCATCTTGCTATC	8636
QY	3215	TATGTATCACACTGTGCATCAGAGAGACCCAGAAAATACCAAAATGCTGGCTGACCGA	3274
Db	8637	TATGTATCACACTGTGCATCAGAGAGACCCAGAAAATACCAAAATGCTGGCTGACCGA	8696
QY	3275	GAACTACCTGCAAGAAATCACAATTCCTGGGAAACAACCTGTTACACAAAGCATCCCTG	3334
Db	8697	GAACTACCTGCAAGAAATCACAATTCCTGGGAAACAACCTGTTACACAAAGCATCCCTG	8756
QY	3335	TGTGAGTGGGAGTACCTGCAATCTGACACAAATGCTCAGAAACAGATCAGTGTCTTAGA	3394
Db	8757	TGTGAGTGGGAGTACCTGCAATCTGACACAAATGCTCAGAAACAGATCAGTGTCTTAGA	8816
QY	3395	GACTCCACTCTTTCTTCCACTTCCAAAGCATGAGGCTCATTTCTGAAGCAGACCAACTGA	3454
Db	8817	GACTCCACTCTTTCTTCCACTTCCAAAGCATGAGGCTCATTTCTGAAGCAGACCAACTGA	8876
QY	3455	GCAAAACCTGTGTCCGGCAGTGTCAACATGGTAAATGGCCAGAGTTATCGAGGCACATT	3514
Db	8877	GCAAAACCTGTGTCCGGCAGTGTCAACATGGTAAATGGCCAGAGTTATCGAGGCACATT	8936
QY	3515	CTCCACCACTGTCAACAGAGGACATGTCAATTTGATGATCCATGACACCAACCGGCA	3574
Db	8937	CTCCACCACTGTCAACAGAGGACATGTCAATTTGATGATCCATGACACCAACCGGCA	8996
QY	3575	TCAGAGAGACCCAGAAAACTACCAAAATGATGGCTGACAAATGAATCTACTGCAGAAATCC	3634
Db	8997	TCAGAGAGACCCAGAAAACTACCAAAATGATGGCTGACAAATGAATCTACTGCAGAAATCC	9056

Oy	3635	AGATGCCGATACAGAGCCCTTGGTGTGTTTACCATGAGACCCCAAGATCAAGGTGGGAATACAG	3694
Dp	9057	AGATGCCGATACAGAGCCCTTGGTGTGTTTACCATGAGACCCCAAGATCAAGGTGGGAATACAG	9116
Oy	3695	CAACCTGAGCGGATGCTCAGACACAGAAAGGACGTGGTGCCTCTCCAGCTCATCTCA	3754
Dp	9117	CAACCTGAGCGGATGCTCAGACACAGAAAGGACGTGGTGCCTCTCTCCAGCTCATCTCA	9176
Oy	3755	GGTTTCCAAAGCCTTAGGGCCCTCTCTTCCGAACAAGACTGTATGTGTTGGAAATGGGAAAGATA	3814
Dp	9177	GGTTTCCAAAGCCTTAGGGCCCTCTCTTCCGAACAAGACTGTATGTGTTGGAAATGGGAAAGATA	9236
Oy	3815	CCGGGGCAGAGAGGACACCACTGTTACTGTGGACGCCATGCCAGATGGCCCTCCACGA	3874
Dp	9237	CCGGGGCAGAGAGGACACCACTGTTACTGTGGACGCCATGCCAGAAATGGGCTCCACGA	9296
Oy	3875	GGCCCATAGACAGACGCTTATTCAGGGACAAATTAATGGGCGAGCTGTGGAAAAAAA	3934
Dp	9297	GGCCCATAGACAGACGCTTATTCAGGGACAAATTAATGGGCGAGCTGTGGAAAAAAA	9356
Oy	3935	TTATCGCCGTAACCTCGATGAGTGTACATCAATAGTCCCGTGGTGTACACAAATGATTCAG	3994
Dp	9357	TTATCGCCGTAACCTCGATGAGTGTACATCAATAGTCCCGTGGTGTACACAAATGATTCAG	9416
Oy	3995	AAAACTTTTGACTACTGTGATATCCCTCTCTGTGCATCCTCTTCAATTTGATGTGGGAA	4054
Dp	9417	AAAACTTTTGACTACTGTGATATCCCTCTCTGTGCATCCTCTTCAATTTGATGTGGGAA	9476
Oy	4055	GGCTCAAGTGGACCCGACAAATAATGTCTCTGGAAAGATTTGAAAGGGGTGTGGCCACCC	4114
Dp	9477	GGCTCAAGTGGACCCGACAAATAATGTCTCTGGAAAGATTTGAAAGGGGTGTGGCCACCC	9536
Oy	4115	ACATTCCTGGCCCTGGCAAGTCAAGTCTCAGAACAGAGTTTGGAAAGACCTTCTGTGGAG	4174
Dp	9537	ACATTCCTGGCCCTGGCAAGTCAAGTCTCAGAACAGAGTTTGGAAAGACCTTCTGTGGAG	9596
Oy	4175	CACCTTAATATCCCCACAGAGTGGGTGTGACTGCTGCTCACTGCTTGAAGAATCCTCAAG	4234
Dp	9597	CACCTTAATATCCCCACAGAGTGGGTGTGACTGCTGCTCACTGCTTGAAGAATCCTCAAG	9656
Oy	4235	GGCTTACCTCAAAAGGTATCTGTGGGTGCACACCAAAATGAAACCTCCAAATCTCATGT	4294
Dp	9657	GGCTTACCTCAAAAGGTATCTGTGGGTGCACACCAAAATGAAACCTCCAAATCTCATGT	9716
Oy	4295	TCAGGAATATGAAGTGTCTAGAGCGCTGTTCTTTGGAGCCCAACAACAGATATTTGGCTTGCT	4354
Dp	9717	TCAGGAATATGAAGTGTCTAGAGCGCTGTTCTTTGGAGCCCAACAACAGATATTTGGCTTGCT	9776
Oy	4355	AAAGCTAAGCAGGCGCTGCCTCATCACTACGACAAAGTAATGGCAGCTTGTCTGCCATCCC	4414
Dp	9777	AAAGCTAAGCAGGCGCTGCCTCATCACTACGACAAAGTAATGGCAGCTTGTCTGCCATCCC	9836
Oy	4415	AGACTACATGTGTACCGCCACAGACTAATGTTATCATCTACGTGGCTGGGAGAAACCCAGG	4474
Dp	9837	AGACTACATGTGTACCGCCACAGACTAATGTTATCATCTACGTGGGAGAAACCCAGG	9896
Oy	4475	TACCTTTGGGACATGGCCCTTCTCAAGGAACCCAGCTCCTGTTATTTGGAATGAAGTGTG	4534
Dp	9897	TACCTTTGGGACATGGCCCTTCTCAAGGAACCCAGCTCCTGTTATTTGGAATGAAGTGTG	9956
Oy	4535	CAATTCATTAATAGTATATTTTGTGCTGAGCATTTTGGCCAGAGGCATGCAAGTTCACAGG	4594
Dp	9957	CAATTCATTAATAGTATATTTTGTGCTGAGCATTTTGGCCAGAGGCATGCAAGTTCACAGG	10011
Oy	4595	TGACAGTGGAGAGCGCTGTGGTTTGTCTTCAGAAAGAACAAATTAATNTTACAAAGAGACAC	4654
Dp	10017	TGACAGTGGAGAGCGCTGTGGTTTGTCTTCAGAAAGAACAAATTAATNTTACAAAGAGACAC	10077
Oy	4655	TTCTTTGGGGTCTTGGCTGTCAAGCCCAATTAAGCGCTGGTGTCTATGCTGTCTGTTTAAG	4714
Dp	10077	TTCTTTGGGGTCTTGGCTGTCAAGCCCAATTAAGCGCTGGTGTCTATGCTGTCTGTTTAAG	10133
Oy	4715	GTTTGTACTTGTGAATGAGGAATGATGAGAAATTAATTAATTTGACGGGAGACAGATGA	4774

Accession	Sequence	Position
D8	10137 GTTTGTTACTTGGATTTGAGGGAATGATGAAATAATTAATTATGGACGGAGACAGAGTGA	10196
OY	4775 AGCATCAACTACTTGTGAAGCTGAAAACGTGGGTAAAGATTTTAGCATGCTGGGAATTAATAG	4834
D8	10197 AGCATCAACTACTTGTGAAGCTGAAAACGTGGGTAAAGATTTTAGCATGCTGGGAATTAATAG	10256
OY	4835 ACACCAATCAAAAGAGACACTGTGTTCCAGCTACACAGCTATGCCAAACCTTGGCATTTTT	4894
D8	10257 ACACCAATCAAAAGAGACACTGTGTTCCAGCTACACAGCTATGCCAAACCTTGGCATTTTT	10316
OY	4895 GGTATTTTGTGTATAGCTTTTAAGTGCTGACGTGACAAATTCGTATTAAGGTGCATA	4954
D8	10317 GGTATTTTGTGTATAGCTTTTAAGTGCTGACGTGACAAATTCGTATTAAGGTGCATA	10376
OY	4955 GCTATGACATTTGTAAAAATAAAGCTGTGCACCTATTTTGTATTTGA 5000	
D8	10377 GCTATGACATTTGTAAAAATAAAGCTGTGCACCTATTTTGTATTTGA 10422	

	RESULT 2	
	US-09-946-893-1	
	/ Sequence 1, Application US/09946893	
	/ Patent No. US20020072494A1	
	/ GENERAL INFORMATION:	
	/ APPLICANT: Cao, Yihai	
	/ TITLE OF INVENTION: Materials and methods relating to endothelial cell growth	
	/ TITLE OF INVENTION: Inhibitors	
	/ FILE REFERENCE: Mewburn	
	/ CURRENT APPLICATION NUMBER: US/09/946, 893	
	/ CURRENT FILING DATE: 2001-09-05	
	/ PRIOR APPLICATION NUMBER: US 60/230,893	
	/ PRIOR FILING DATE: 2000-09-05	
	/ NUMBER OF SEQ ID NOS: 11	
	/ SOFTWARE: Patentin Ver. 2.1	
	/ SEQ ID NO 1	
	/ LENGTH: 2497	
	/ TYPE: DNA	
	/ ORGANISM: Homo sapiens	
	/ FEATURE:	
	/ NAME/KEY: CDS	
	/ LOCATION: (50)..(2482)	
	US-09-946-893-1	
QY	Query Match	22.1%; Score 1106.4; DB 10; Length 2497;
Bt	Best Local Similarity	81.4%; Pred. No. 0;
Mtches	Matches 1334; Conservative	0; Mismatches 241; Indels 63; Gaps 2;
Dy	3159 TATCGAGGCAATTCCTCCACCACTGTCTCAGGAAGAAGACTGTCAATCTGGTCACTTAG	3218
Dy	896 TATCGAGGGAATGTGGTGTTACGCTGTCCGGGCACAACGTGTGACACTGAGTGCAC	955
Dy	3219 ATACCACTGCGCATCAAGAGAACCCAGAAAACCTACCCTAAATGCTGGCCCTGACCGAGAAC	3278
Dy	956 ACCCCTCACACATTAACAGGACACCAAAAACTCCCTCGCAAAAATTGGATGAAC	1015
Dy	3279 TACTGAGGAATCCAGATTTCTGGAAACAACCCCTGGTGTTCACACACCGATTCGGTGTG	3338
Dy	1016 TACTGCCGAATCTCGACGAGAAAAAGGCCCCCATGTGCCATTAACCAACAGCAAGTG	1075
Dy	3339 AGGTGGGAGTACTGCATACTGTGACACATGCTCAGAAACGATCAGGTGCTCTAGAGACT	3398
Dy	1076 CGATGGGAATTACTGTAAAGATACCGTCTG-----	1104
Dy	3399 CCCACTGTGTCCAGCTTCCAAAGCATGTGAGGCTCATTTCTGAAGCAGACCAACTGAGCAA	3458
Dy	1105 -----TGACTCTCCCCAGATCCAGGACAACATTGGCTCCACAGCAACCACTGAGCTA	1159
Dy	3459 ACCCCTGTGTGTCGGCAGTGTACCATGTAATGCGCAGAGTTATTCGAGGCACATTTCTCC	3518
Dy	1160 ACCCCTGTGTGTCAGAGACTGCTACCATGTTGATGSACAGACTTACGAGGACATCTCTCC	1219
Dy	3519 ACCACTGTCAACGAGGAAGCATGTCAATCTTGTTGTCATCATGACACCAACCGGATCAG	3578

```
Db 1220 ACCACACACAGAGAAAGTGTACAGTCTTGGTCACTTATGACACACACCGGACACG 1279
QY AGGACCCAGAAAATCTACCCAAATGATGGCTGACAAATGAATCTACTGAGAAATCCAGAT 3638
Db 1280 AAGACCCAGAAAATCTACCCAAATGATGGCTGACAAATGAATCTACTGAGAAATCCAGAT 1339
QY GCGATACAGGCGCTTGGTGTATTTACATGAGACCCAGCAATCAAGTGGGAGTACTGCAAC 3698
Db 1340 GCGATACAGGCGCTTGGTGTATTTACATGAGACCCAGCAATCAAGTGGGAGTACTGCAAC 1399
QY CTGAGCGGATGCTCAGACAGAGAGGAGTGGTCCGCTCCGATCTGATCTCAGGTT 3758
Db 1400 CTGAAAAATGCTCAGAGAGAGAGAGGAGTGGTGGTACACCTCGCGCTGTGTCTGCTT 1459
QY 3759 CCAAGCTAGGGCTCTCTTCTGAAACAAGATGATGTTTGGGAATGGGAAAGATACCGG 3818
Db 1460 CCAAGATGTAGAGACTCTCTCCGAAAGAGATGATGTTTGGGAATGGGAAAGATACCGA 1519
QY 3819 GGCAGAGAGGACCACTGTTTACTGAGAGCCATGCGCAGGAATGGGCTGCCAGAGGCC 3878
Db 1520 GGCAGAGAGGAGCACTGTTTACTGAGAGCCATGCGCAGGAATGGGCTGCCAGAGGCC 1579
QY 3879 CATGACACAGCAGCTTCAATTCAGGAGCAAAATGAATGGCAGGCTTGAAAAAATTAC 3938
Db 1580 CATGACACAGCAGCTTCAATTCAGGAGCAAAATGAATGGCAGGCTTGAAAAAATTAC 1639
QY 3939 TGCGGTAAACCTGATGTGATGATCAATGATGCTGCTGATCAACAAATGATGCAAGAAA 3998
Db 1640 TGCGGTAAACCTGATGTGATGATGATGATGCTGCTGATCAACAAATGATGCAAGAAA 1699
QY 3999 CTTTGTACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4058
Db 1700 CTTTGTACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1759
QY 4059 CAATGAGAGCGGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4118
Db 1760 CAATGAGAGCGGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1819
QY 4119 TCCTGGGCTCTGGCAAGTCACTGTCAGAAACAAGTTTGGAAAGCACTGCTGAGAGGACC 4178
Db 1820 TCCTGGGCTCTGGCAAGTCACTGTCAGAAACAAGTTTGGAAATGCACTCTGAGAGGACC 1879
QY 4179 TTAATATCCCAAGAGTGGTGTGACTGCTGATGATGATGATGATGATGATGATGATGAT 4238
Db 1880 TTAATATCCCAAGAGTGGTGTGACTGCTGATGATGATGATGATGATGATGATGATGAT 1939
QY 4239 TCATCCACAGAGTCACTGCTGGGTGACACACCAAGAGTGAATCCGATGTCATGTTCCAG 4298
Db 1940 TCATCCACAGAGTCACTGCTGGGTGACACACCAAGAGTGAATCCGATGTCATGTTCCAG 1999
QY 4299 GAAATAGAGTGTCTAGAGTGTCTTGGAGCCACACACAGATATGTCCTGCTAAAG 4358
Db 2000 GAAATAGAGTGTCTAGAGTGTCTTGGAGCCACACACAGATATGTCCTGCTAAAG 2059
QY 4359 CTAAAGCAGGCTGCTGATCACTGACAAAGTAATGAGCAGCTTGTGTCATCCAGAGC 4418
Db 2060 CTAAAGCAGGCTGCTGATCACTGACAAAGTAATGAGCAGCTTGTGTCATCCAGAGC 2119
QY 4419 TACATGCTACCGCCAGAGTGAATGTTACATGATGCTGGGAGAGAAACCCAGAGTACC 4478
Db 2120 TACATGCTACCGCCAGAGTGAATGTTACATGATGCTGGGAGAGAAACCCAGAGTACC 2179
QY 4479 TTTGGAGTGGCTTCTCAAGAGAGCCAGAGCTCTGTTATTTAGATGATGATGATGATGAT 4538
Db 2180 TTTGGAGTGGCTTCTCAAGAGAGCCAGAGCTCTGTTATTTAGATGATGATGATGATGAT 2239
QY 4539 CACTATAAGT-----AATTTGCTGCTGAGCAATTTGGCC 4571
Db 2240 CCGATATAGTCTTGAATGGAAGATCCAAATCCAGCAACTGTGCTGGGCAATTTGGCC 2299
QY 4572 AGAGGCACTGACAGTTGCCAGGGTGACAGTGGAGGGCTCTGTTGCTTCCAGAAAGGAC 4631
```

```
Db 2300 GGAAGCATTGACAGATTTGCCAGGGTGACAGTGGAGTCTCTGTGTTGCTTGACAGAGAC 2359
QY 4632 AATAATATTTTACAGAGATCACTTCTTGGGCTTGTGGCTGTGACAGCCCAATTAAGCTT 4691
Db 2360 AATAATATTTTACAGAGATCACTTCTTGGGCTTGTGGCTGTGACAGCCCAATTAAGCTT 2419
QY 4692 GGTGTCTATGCTGCTGCTTCAAGTGTGTTGTTACTTGATTTGAGGAGAAATATATAT 4751
Db 2420 GGTGTCTATGCTGCTGCTTCAAGTGTGTTGTTACTTGATTTGAGGAGTATGAAAAATAT 2479
QY 4752 TAATTGACGGAGACAG 4769
Db 2480 TAATTGACGGAGACAG 2497
```

RESULT 3
US-09-924-340-53
Sequence 53, Application US/09924340
Publication No. US20030027248A1

GENERAL INFORMATION:
APPLICANT: Bejatin, Stephane
APPLICANT: Tanaka, Hiroaki
TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
FILE REFERENCE: 91, US2, REG
CURRENT APPLICATION NUMBER: US/09/924, 340
CURRENT FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: US 60/305,456
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 60/302,277
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/298,698
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 60/293,574
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 112
SOFTWARE: JPatent
SEQ ID NO 53
LENGTH: 1907
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..1043
NAME/KEY: CDS
LOCATION: 1044..1664
NAME/KEY: 3'UTR
LOCATION: 1665..1907
NAME/KEY: POLYA_Signal
LOCATION: 1869..1874
NAME/KEY: POLYA_site
LOCATION: 1892..1907

US-09-924-340-53

Query Match 13.1%; Score 654.6; DB 9; Length 1907;
Best Local Similarity 86.2%; Pred. No. 5e-196;
Matches 766; Conservative 0; Mismatches 79; Indels 44; Gaps 2;

```
QY 4132 AAGTAGCTGTGAGAACAGGTTTGGAAAGCACTTGTGAGAGCACTTAATATCCAG 4191
Db 1018 AACTTAATCTCTTTCCAGGTTTGGAAATGCACTTGTGAGAGCACTTGAATATCCAG 1077
QY 4192 AGTGGCTGCTGACTGCTCTCACTGCTTGAAGAAGTCTCTCAAGCCTTCAATCTTCAAGG 4251
Db 1078 AGTGGCTGCTGACTGCTCTCACTGCTTGAAGAAGTCTCTCAAGCCTTCAATCTTCAAGG 1137
QY 4252 TCATCTCTGGGTGACACCAAGAGTGAATCTGATCTCATGTTCCAGAAATAGAGTGT 4311
Db 1138 TCATCTCTGGGTGACACCAAGAGTGAATCTGATCTCATGTTCCAGAAATAGAGTGT 1197
QY 4312 CTAGGCTGTCTTGGAGGCCACACAGCAGATATTTGCTGCTTAAGCTAAGCAGGCGTG 4371
Db 1198 CTAGGCTGTCTTGGAGGCCACACAGCAGATATTTGCTGCTTAAGCTAAGCAGGCGTG 1257
```

OY	4372	CCGCAATCAGTGCACAAAGTAATGCCAGCTGTGTCGCATCCCGACATCACATACATGGTCAACG	4433
Db	1258	CCGTCATCATCTGCACAAGTAATCCACACTTTGTCTGCCATCCCCAATTATATGTGTCGCTG	1317
OY	4432	CACAGACATGAATTTACATCATCTAGCTGGTGCGGAGAAAACCCAAAGTACCCTTTGGACTGGCC	4491
Db	1318	ACCGGACCGGAATGTTTCATCATCTAGCTGGTGCGGAGAAAACCCAAAGTACCCTTTGGAGCTGGCC	1377
OY	4492	TTCCTCAAGAACCCCACTCTCTGTTATTGAATAGAATGTGTGCATACTACTATACT---	4548
Db	1378	TTCCTCAAGAACCCCACTCTCTGTTATTGAATAGAATGTGTGCATACTACTATACTGTTTC	1437
OY	4549	-----ATATTTTGTGCTGAGACATTTGGCCAGAGGACTGTACA	4584
Db	1438	TGAATGGAAGAAGTCCATCTCCACCGAACTGTGTCTGTGGCATTTGGCCGAGAGCACTGACA	1497
OY	4585	GTTGCCAGGCTGACAGTGGAGGCTCTGTGTTTTGCTTGAGAGAGCAAAATPACATTTTAC	4644
Db	1498	GTTGCCAGGCTGACAGTGGAGGCTCTGTGTTTTGCTTGAGAGAGCAAAATPACATTTTAC	1557
OY	4645	AAGGAGCACTCTTGTGGGCTGTGGCTGTGCAAGCCCAATTAAGCTGGTGTCTATGGCTC	4704
Db	1558	AAGGAGCACTCTTGTGGGCTGTGGCTGTGCAAGCCCAATTAAGCTGGTGTCTATGGCTC	1617
OY	4705	GTTGTTCAAGTTTGTACTTCTTGATTTGAGGAGATGATGAAAAATTAATTAATTGACGGGA	4764
Db	1618	GTTGTTCAAGTTTGTACTTCTTGATTTGAGGAGATGATGAAAAATTAATTAATTGACGGGA	1677
OY	4765	GACAGATGAGACATCAACTACTTTGAAAGCTGAAACGTGGGTAAAGATTATAGCATGCTG	4824
Db	1678	GACAGATGAGACATCAACTACTTTGAAAGCTGAAACGTGGGTAAAGATTATAGCATGCTG	1737
OY	4825	GAATTAATTAACACAGCATCAACAGAGACACTGTTCCACAGTACCAGTATGCCAAACT	4884
Db	1738	GAATTAATGCGAGTATCAAAGAGACACTGTCCCGAGTACCAGTATGCCAAACT	1797
OY	4885	TGCGATTTTGGTATTTTGTGTATAAGCTTTTAAGGTGTGACTGCACAAATTCGTATTA	4944
Db	1798	CGCGATTTTGTGTATTT-----TTCTGACTGTGATCTGTGTAGTA	1840
OY	4945	AGGTGATAGTATGATGACATTTGTTAAAAATAAATCTGCACATTTATTTT	4993
Db	1841	AGGTGATAGTATGATGACATTTGTTAAAAATAAATCTGTACTTACTTAACTT	1889
 RESULT 4 US-09-992-600A-53 Sequence 53, Application US/09992600A Publication No. US20030027161A1 GENERAL INFORMATION: APPLICANT: Benjamin, Stephanie FILE REFERENCE: Tanaka, Hiroaki TITLE OF INVENTION: HUMAN CDNAs AND PROTEINS AND USES THEREOF CURRENT APPLICATION NUMBER: US/09/992, 600A CURRENT FILING DATE: 2001-07-13 PRIOR APPLICATION NUMBER: US 09/924,340 PRIOR FILING DATE: 2001-08-06 PRIOR APPLICATION NUMBER: PCT/IB01/01715 PRIOR FILING DATE: 2001-08-06 PRIOR APPLICATION NUMBER: US 60/305,456 PRIOR FILING DATE: 2001-07-13 PRIOR APPLICATION NUMBER: US 60/302,277 PRIOR FILING DATE: 2001-06-29 PRIOR APPLICATION NUMBER: US 60/298,698 PRIOR FILING DATE: 2001-06-15 PRIOR APPLICATION NUMBER: US 60/293,574 NUMBER OF SEQ ID NOS: 114 SOFTWARE: Jpatent SEQ ID NO 53 LENGTH: 1907 TYPE: DNA			

	ORGANISM	Homo sapiens	
:	FEATURE:		
:	NAME/KEY:	5'UTR	
:	LOCATION:	1..1043	
:	FEATURE:		
:	NAME/KEY:	CDS	
:	LOCATION:	104..1664	
:	FEATURE:		
:	NAME/KEY:	3'UTR	
:	LOCATION:	1665..1907	
:	FEATURE:		
:	NAME/KEY:	polyA_signal	
:	LOCATION:	1869..1874	
:	FEATURE:		
:	NAME/KEY:	polysite	
:	LOCATION:	1892..1907	
:			
:US-09-992-	-600A-53		
	Query Match	13.1%; Score 654.6;	DB 9; Length 1907;
	Best Local Similarity	86.2%; Pred. No.	5e-196;
Matches	766; Conservative	0; Mismatches	79; Indels 44; Gaps 2
OY	4132	AAGTCATCTCACAACAAAGTTTGGAAAAGCACTTCTGTGGAAGCGAACCTTAATTATCCCCAG	4191
Dd	1018	AACTAAATCCCTCTTCAGAGTTTGAAATCACCTTCGTGTGGAGGCACCCTGTATCCCAG	1077
OY	4192	AGTGGGTGTGACTGCTGCTCATCTGCTTGAANAAGTCCCTCAAAGGCCCTTCATCTACAG	4251
Dd	1078	AGTGGGTGTGACTGCTGCTGCCACTGCTTGGAAAACTCCCCAAGGCTTCATCTACAGG	1137
OY	4252	TCACTCTGGGTGCACACCAAGAATGACCTGCAATCTCATGTTCAGAAATAGAACTGT	4311
Dd	1138	TCACTCTGGGTGCACACCAAGAATGAAATCTGAAACCCAGTAGTTCAGAAATAAGATGT	1197
OY	4312	CTAGAGCTGTCTTGGAGGCCACACAAAGCATATTGCCCTGTGAAGCTAAGAGGCGTG	4371
Dd	1198	CTAGAGCTGTCTTGGAGGCCACACGAAAGATATTGCTTGTGAAGCTAAGCAAGCTCTG	1257
OY	4372	CCGTCACTACTGACAAAGATATGCCCCAGCTTGTCTGCCATCCCCAGACTATAGTGCACG	4431
Dd	1258	CCGTCACTACTGACAAAGATATGCCAGCTTGTCTGCCATCCCCAATATATGTGGTGCCTG	1317
OY	4432	CCAAGCATGATGTTCATCATCTGGGCGGGAGAAACCAAGATACCTTGGAGCTGGCC	4491
Dd	1318	ACCGACCGCATGTTCATCATCTGGGCGGGAGAAACCAAGATATCTTGGAGCTGGCC	1377
OY	4492	TTCCTCAAGAAAGCCACAGCTCTGTATTGGAATAAGATGTGCATCATATATAGT---	4548
Dd	1378	TTCCTCAAGAAAGCCACAGCTCTGTATTGGAATAAGATGTGCATCATATATAGT---	1437
OY	4549	-----AATTTGTGCTGAGCATTTTGGCCACAGAGGCACTGCACA	4584
Dd	1438	TGAATGGAAGATCCATCCACCGAACTGTGTCTGGGCAATTTGGCCGAGGCACTGCACA	1497
OY	4585	GTTGCCAGGAGGACAGTGGAGGCGCTCTGTTTTGCTTCAGAAAGACAAATACATTTTAC	4644
Dd	1498	GTTGCCAGGAGGACAGTGGAGGCGCTCTGTTTTGCTTCAGAAAGACAAATACATTTTAC	1557
OY	4645	AAGGATCACTCTTGGSGGTCTTGGCTGTGCACGCCGCCCAATATAGCTGTCTATGCTC	4704
Dd	1558	AAGGATCACTCTTGGSGGTCTTGGCTGTGCACGCCGCCCAATATAGCTGTCTATGCTC	1617
OY	4705	GTTGTTCAAGTTGTTGACTTGGATTTGAGGGAATGATGGAATAATTAATTGGACGGGA	4764
Dd	1618	GTTGTTCAAGTTGTTGACTTGGATTTGAGGGAATGATGGAATAATTAATTGGACGGGA	1677
OY	4765	GACAGAGTGAAAGTCATCAACCTACTTTGGAAGCTGAAACGTGGGTAAAGATTTAGCATGCT	4824
Dd	1678	GACAGAGTGAAAGTCATCAACCTACTTTGGAAGCTGAAACGTGGGTAAAGATTTAGCATGCT	1737
OY	4825	GAATATATAGACAGCAATCAAAGAAAGACAGTGTTCACAGCTACAGCATATGCCAAACT	4884
Dd	1738	GAATATATAGCTATATCAAAGAAAGACAGTGTTCACAGCTACAGCATATGCCAAACT	1797

QY	1841	4945	1798	QY	4885
AGGTGCATACGCTATGACACTTTGTTAAATAAATACTGCACTTATTTT	AGGTGCATACGCTATGACACTTTGTTAAATAAATACTGCTACTTACTT	AGGTGCATACGCTATGACACTTTGTTAAATAAATACTGCACTTATTTT	CGGCATTTTTTGCTATT	TGCATTTTGGATTTTGTGATTAACGTTTAAAGTGCAGTACAAATCTGTATTA	TGCATTTTGGATTTTGTGATTAACGTTTAAAGTGCAGTACAAATCTGTATTA
1889	1893	4993	1840	4944	4944

```

RESULT 5
; US-09-880-107-2341
; Sequence 2341, Application US/09880107
; Patient No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-US
; CURRENT APPLICATION NUMBER: US/09/880, 107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2341
; LENGTH: 1236
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M34276
; US-09-880-107-2341

```

Query Match	Similarity	6.1%	Score 306.4	DB 10	Length 1236
Best local	Similarity	87.0%	Pred. No. 5,4e-86		
Matches 355	Conservative	0	Mismatches 36	Indels 17	Gaps 1
QY 4591	AGGGTACAGTGGAGGGCCCTGTGGTTGCTTCGAGAAAGACAAATATACAAATTTTACAMGAG	4650			
Db 39	AGGGTACAGTGGAGGGCCCTGTGGTTGCTTCGAGAAAGACAAATATACATTTTACAAGGAG	98			
QY 4651	TCACCTTGGGGGCTGTGGCTGTGCAGCCCCCAATAAAGCCGTGCTATGCTCGTGT	4710			
Db 99	TCACCTTGGGGGCTGTGGCTGTGCAGCCCCCAATAAAGCCGTGCTATGCTCGTGT	158			
QY 4711	CAAGTTTCTTACTTGATGGATTGAGGAAATGATGAGAAATTAATTAATTGAGAGGAGACAG	4770			
Db 159	CAAGTTTCTTACTTGATGGATTGAGGAGATGATGAGAAATTAATTAATTGAGAGGAGACAG	218			
QY 4771	GTGACACATCAACCTACTTACAGACTGAACGTGGGTGAGGATTTAGCATGCTGGAAATA	4830			
Db 219	GTGACACATCTGACTACCTACCTAGAGGCTGGAGAGTGGGTAGGATTTACATGCTGGAAATA	278			
QY 4831	ATGACAGCAATCAACGAGACACTGTGCCAGTACACAGCTGTGCAACCTGTGGCAAT	4890			
Db 279	ACTGGCAATCAACGAGACACTGTGCCAGTACACAGCTGTGCAACCTGTGGCAAT	338			
QY 4891	TTTTGGTATTTTGTGTATTAAGCTTTTAAGTCTGACTGACAAATTTCTGTATTAAGTGT	4950			
Db 339	TTTTTGTGTATTTTGTGTATTAAGCTTTTAAGTCTGACTGACAAATTTCTGTATTAAGTGT	381			
QY 4951	CATAGCTATGACATTTGTTAAAAATTAACCTGCACCTATTTGATTT	4998			
Db 382	CATAGCTATGACATTTGTTAAAAATTAACCTGCACCTATTTGATTT	429			

RESULT 6
US-09-873-676-29
; Sequence 29, Application US/09873676

```

? Patent No. US20020077289A1
? GENERAL INFORMATION:
? APPLICANT: Macdonald, Nicholas J.
? APPLICANT: Sim, Kim L.
? TITLE OR INVENTION: Angiostatin and Endostatin Binding Proteins and Methods of Use
? FILE REFERENCE: 05213-0378 (43170-259333)
? CURRENT APPLICATION NUMBER: US/09/873,676
? CURRENT FILING DATE: 2001-06-04
? PRIOR APPLICATION NUMBER: US 60/209,065
? PRIOR FILING DATE: 2000-06-02
? PRIOR APPLICATION NUMBER: US 60/289,387
? PRIOR FILING DATE: 2001-05-08
? NUMBER OF SEQ ID NOS: 123
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 29
? LENGTH: 1134
? TYPE: DNA
? ORGANISM: Homo sapiens
? IS-09-873-676-29

```

Query Match	Similarity	Score	DB	Length
Best local	56.81%	257.2	10	1134
Matches	401	Conservative	0	Mismatches 163; Indels 36; Gaps 1;

RESULT 7
US-09-960-352-10993
; Sequence 10993, Application US/09960352
; Patent No. US2002013719A1
; GENERAL INFORMATION
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tac, Nengbing
; APPLICANT: Byalt, John C.

```

; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 10993
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (191)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 47-LIB34-003-Q1-E1-D12
US-09-960-352-10993

Query Match          4.7%; Score 235.6; DB 10; Length 384;
Best Local Similarity 77.3%; Pred. No. 5.8e-64;
Matches 299; Conservative 0; Mismatches 85; Indels 3; Gaps 1;

QY 4067 GCCGAGAAATGTCCTGGAGACATTGTAGGGGGGTGTGGCCCAACCCACATTCCTGGCC 4126
    |||||
DB 1 GCCGAGAAATGTCCTGGAGACATTGTAGGGGTGTGTGCCAACCCACACTCTTGGCC 60

QY 4127 CTGGCAAGTCAGTCTGCAGACAGAGTTTGGAAAGCACTTCTGTGGAGGACCTTAATATC 4186
    |||||
DB 61 CTGGCAAGTCAGCTTTCGAGAAAGCT---CGAGGCACTTCTGTGGAGAACTGTGATATC 117

QY 4187 CCCAGAGTGGTGTCTGAGCTGTCTGCTGCTTGAAGAGTCTTCAAGGCTTATCTCA 4246
    |||||
DB 118 CCCAAGTGGTGTCTGAGCTGTCTGCTGCTTGAAGAGTCTTGAAGCTTATCTCA 177

QY 4247 CAAGGTATCTGTGGTGCACACAGAGAGTGAACCTCGATCTCAAGCTTCAAGAAATAGA 4306
    |||||
DB 178 CAAGGTATCTGTGGTGCACACAGAGAGTGAACCTCGAGGAGAGTGTCCAGGAAATACC 237

QY 4307 AGTGTAGAGCTGTTCTTGGAGGCCACACAGAGATATTGCTTCTAAAGCTAAGCAG 4366
    |||||
DB 238 AGTGTCAAGCTGTTCGGGAGGCCCTCTCAGGGGAGCATTTGCTTCTCAAGCTAGAGCA 297

QY 4367 GCCTGGCGGTATCATCTACAAAGTAATGCGAGCTTGTCTGCCATCCCGAGACTACATGT 4426
    |||||
DB 298 ACCGCGCATCATCAGCAAGAGAGTAATCCAGCTGTGTGCCACCCCGAATCATGTGT 357

QY 4427 CACCGCAGAGCTGAATGTATCATCAC 4453
    |||||
DB 358 TGCAGCCCGGAGCAATGTATCATCAC 384

RESULT 8
US-09-960-352-10851
; Sequence 10851, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengping
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 10851
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 46-LIB34-086-Q1-E1-D6
US-09-960-352-10851
```

```

Query Match          4.6%; Score 231.4; DB 10; Length 468;
Best Local Similarity 72.6%; Pred. No. 1.4e-62;
Matches 326; Conservative 0; Mismatches 96; Indels 27; Gaps 1;

QY 4223 GAAGTCTCAAGGCTTCTATCTTACAGAGTCTATCTGTGGTGCACACCAAGTGAACCT 4282
    |||||
DB 20 GAACATTTTACGGCTGTCTATCTTACAGAGTCTATCTGTGGTGCACACCAAGTGAACCT 79

QY 4283 CGAATCTATGTTCAGAGAAATAGAAGTGTCTAGGCTGTCTTGGAGCCACACAGAGCA 4342
    |||||
DB 80 GGAACAGAGTGTCCAGAAATACCAAGTGTCTAGGCTGTCTTGGAGCCACACAGAGCA 139

QY 4343 TATTGCTTGTAAAGCTAAGCAGGCTGCGGTATCATCTGACAAATATATGCAAGCTTG 4402
    |||||
DB 140 CATTGCTTATCTCAAGCTGACAGACCTGCTCATATCAGAAAGAGTATATCCAGCTTG 199

QY 4403 TCTGCAATCCCAAGACTACATGATGTCACCGCAGAGCTGAATGTATATCATCTGCTGGG 4462
    |||||
DB 200 TCTGCAATCCCAAGACTACATGATGTCACCGCAGAGCTGAATGTATATCATCTGCTGGG 259

QY 4463 AGAAGCCCAAGTACCTTGTGGAGCTGCTTCAAGAGAGCCAGCTTGTATTTGA 4522
    |||||
DB 260 AGAAGCCCAAGTACCTTGTGGAGCTGCTTCAAGAGAGCCAGCTTGTATTTGA 319

QY 4523 GAATGAAGTGTCAATCATCTAATATAT-----ATTGG 4555
    |||||
DB 320 GAACAAAGTGTATATTCACAGAGTATCTGACGGAGGGTCAAGCCACAGAGCTGTG 379

QY 4556 TGCTAGAGATTTGGCAAGGCACTGACATGTCGACAGGTGTACAGTGAAGGCTCTGTGT 4615
    |||||
DB 380 TGCCGGCATGTGATTTGAGGAGCTGACAGCTGTCAGAGGTGACGCCGGCTCTGTGT 439

QY 4616 TTGCTTCGAAGAGCAATATCATTTTAC 4644
    |||||
DB 440 CTGCTTTGAGAAAGCAATATCATTTCTGC 468

RESULT 9
US-09-960-352-7521
; Sequence 7521, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengping
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7521
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 32-LIB34-042-Q1-E1-H7
US-09-960-352-7521

Query Match          4.6%; Score 229.2; DB 10; Length 372;
Best Local Similarity 79.1%; Pred. No. 6e-62;
Matches 299; Conservative 0; Mismatches 73; Indels 6; Gaps 2;

QY 3827 GGCAACCACTGTTACTGGAGCGCATGCGAGAAATGGGCTGCCAGAGAGCCCATAGACA 3886
    |||||
DB 1 GGCAACCAAGGTGGCGGTGTCTGTCAGAGAGTGGCGGCCAGGAGGCCACACACA 60

QY 3887 CAGCAGCTTATCTCCAGGAGCAAAATTAATGGCGAGGTGTGAAAAAATTAATCTCCGTA 3946
    |||||
DB 61 CAGCAGCTTATCTCCAGGAGCAAAACCAAGTCAAGGCTGTGAGAGAACTAATCTCCGTA 120

QY 3947 CCTGATGGTGAATCAATGATGCTGTGCTACATGAATGAAGAAACTTTTGA 4006
    |||||
```

Db 121 TCCTGAGGTGATGTCAACGGCCCTGTGCTACACCATGACCAAGCAAGCTTTTGA 180
QY 4007 CTACTGTGATATCCCTCTCTGTGCATCCTCTTATTTGATGTGGGAAGCCTCAAGTGA 4066
Db 181 CTACTGTGAGTCCCTCAGTGTGAGTCC---TCATTCGACTGTGGAAAGCCCAAGTGA 237
QY 4067 GCCCAAAATGTCCTGGAAGCATTGTAGGGGGGTGTGTGGCCCAACCCCAATTCCTGGCC 4126
Db 238 GCCCAAAATGTTCTGGAAGATGTAGCTGGGTGTGTCCAGCCACACTTGGCC 297
QY 4127 CTGGCAAGTCACTCTCAAGCAAGTTTGGAAACACTTCTGTGAGGCACTTAATATC 4186
Db 298 CTGGCAAGTCAAGCTTCTGCAACAGCT---CGAGGCACTTCTGTGAGGAAGTCTGATATC 354
QY 4187 CCCGAGTGGGTGCTGAC 4204
Db 355 CCCGAGTGGGTGCTGAC 372

RESULT 10

US-09-796-692-5593
; Sequence 5593, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 5593
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-5593

Query Match 4.1%; Score 204.4; DB 9; Length 313;
Best Local Similarity 89.4%; Pred. No. 3.7e-54;
Matches 220; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 3906 ACAATAAATGGGACAGTCTGGAAAAATTTACTGCCGTAAACCCGTGATGTACATCAAT 3965
Db 68 AAAAACCAACACCGCATGTGTCACTCACTAGTACGCGTAACCCGTGATGTACATCAAT 127
QY 3966 GGTCCCTGGTGTACACATGAATCAAGAAACCTTTTGTACTACTGTGATATCCCTGTC 4025

Db 128 GGTCCCTGGTGTACAGATGATCCAAAGAACTTTTGTACTACTGTGATATCCCTGTC 187
QY 4026 TGTGCATCCTCTTCAATTTGATGTGGGAAGCCTCAAGTGGAGCCGAAGAAATGTCTGGA 4085
Db 188 TGTGCATCCTCTTCAATTTGATGTGGGAAGCCTCAAGTGGAGCCGAAGAAATGTCTGGA 247
QY 4086 AGCATTTAGAGGGGTGTGTGGCCCAACATTCCTTGGCCCTGGCAAGTCAAGTCTGAGA 4145
Db 248 ACATTTAGAGGGGTGTGTGGCCCAACATTCCTTGGCCCTGGCAAGTCAAGTCTGAGA 307
QY 4146 ACAAGG 4151
Db 308 ACAAGG 313

RESULT 11

US-09-960-352-11088
; Sequence 11088, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengping
; APPLICANT: Byatt, John C.
; APPLICANT: Matbalaagan, Nagapan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11088
; LENGTH: 394
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 47-LIB34-084-Q1-E1-D4
US-09-960-352-11088

Query Match 3.9%; Score 197; DB 10; Length 394;
Best Local Similarity 72.0%; Pred. No. 9.6e-52;
Matches 283; Conservative 0; Mismatches 85; Indels 25; Gaps 1;

QY 4248 AAGGTATCTGTGGTGCACACCAAGTAAGTGAACCTCGAATCTCATTTTCAGAAATAGAA 4307
Db 2 AAGCATCTGTGGTGCACACCAAGTAAGTGAACCTCGAATCTCATTTTCAGAAATAGAA 61
QY 4308 GTGTCAAGCTGTCTTGGAGCCCAACACCAAGTAATTTGCTTGTCTTAAGCTTAAGCAGG 4367
Db 62 GTGTCAAGCTGTCTTGGAGCCCAACACCAAGTAATTTGCTTGTCTTAAGCTTAAGCAGG 121
QY 4368 CTTGCGGTATCATCTGACAAAGTAATGCAAGCTTGTCTGCATCCCAAGTCAATGCTG 4427
Db 122 CTTGCGGTATCATCTGACAAAGTAATGCAAGCTTGTCTGCATCCCAAGTCAATGCTG 161
QY 4428 ACCGCAAGCACTGAATGTATACATCTGCTGGGGAAGAAACCAAGTACCTTTGGACT 4487
Db 182 GCAGCCGCAAGCACTGAATGTATACATCTGCTGGGGAAGAAACCAAGTACCTTTGGAG 241
QY 4488 GGCCTTCAAGGAAGCCACAGCTCTCTTGTATTGAAATGAAGTCACTCACTAATAG 4547
Db 242 GGCCTTCAAGGAAGCCACAGCTCTCTTGTATTGAAATGAAGTCACTCACTAATAG 301
QY 4548 TATATTGTGCTGAG-----CATTTGGCCAGAGCACTGA 4582
Db 302 TATTCGACGGGAAGTCAAGCCACGAGAGTGTGTGGCCCATCTGATTGAGAGTGA 361
QY 4583 CAGTTGCCAGGCTGACAGTGAAGGCTCTGTGCT 4615
Db 362 CAGTGCACAGTGAAGGAGGAGGCTGTGCT 394

RESULT 12

US-09-960-352-12743


```
/ Sequence 12743, Application US/09960352
/ Patent No. US20020137139A1
/ GENERAL INFORMATION:
/ APPLICANT: Warren, Wesley C.
/ APPLICANT: Tao, Mengding
/ APPLICANT: Byatt, John C.
/ APPLICANT: Mathialagan, Nagappan
/ TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
/ FILE REFERENCE: 16511.006/37-21(10298)C
/ CURRENT APPLICATION NUMBER: US/09/960,352
/ CURRENT FILING DATE: 2001-09-24
/ NUMBER OF SEQ ID NOS: 15112
/ SEQ ID NO 12743
/ LENGTH: 346
/ TYPE: DNA
/ ORGANISM: Bos taurus
/ OTHER INFORMATION: Clone ID: 54-LIB34-056-Q1-E1-F10
US-09-960-352-12743
```

```
Query Match          3.9%; Score 193.8; DB 10; Length 346;
Best Local Similarity 76.8%; Pred. No. 9e-51;
Matches 262; Conservative 0; Mismatches 52; Indels 27; Gaps 1;
```

```
QY 4387 AACGAATGCCAGCTGTCTGCCATCCCAAGACTGATGTCACCGCCAGACTGAATTT 4446
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6 AGGTAATCCAGCTGTCTGCCATCCCAAGACTGATGTCACCGCCAGACTGAATTT 65

QY 4447 ACATCACTGGCTGGGAGAAACCCAGGACCTTGGGAGCTGGCTTCAGGAAAGCCC 4506
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 66 ACATCACTGGCTGGGAGAAACCCAGGACCTTGGGAGCTGGCTTCAGGAAAGCCC 125

QY 4507 AGCTCCCTGTATTGAGATGAAGTGCATCACTATAAGTAT----- 4550
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 126 ACCTGCTGTGATCGAAGCAAGGTGTGTAATCGCAAGATATCGAAGGAGGTGTA 185

QY 4551 -----ATTGTGCTGAGCATTTGGCCAGAGGACCTGACAGTTCGCCAGGGTGACA 4599
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 186 AGCCCAAGGAGCTGTGCTGGCCGATCTGATGAGGAGTACACCTCCAGGGTGACA 245

QY 4600 GTGAGAGGCTGTGTTGCTTCGAGAGACAAATACATTTTACAGGAGTCACTTCTT 4659
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 246 GCGGCGGCGCTGTGCTGTGCTTTGAGAGAGACAAATACATTTTGTGCAAGGAGTCACTTCTT 305

QY 4660 GGGGTCTTGCTGTGCACGCCCCCAATAGCTGTGTCTAT 4700
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 306 GGGCTCTTGCTGTGCACGCCCCCAATAGCTGTGTCTAT 346
```

```
RESULT 13
US-09-833-381-252
/ Sequence 252, Application US/09833381
/ Patent No. US20020132090A1
/ GENERAL INFORMATION:
/ APPLICANT: Robison, Keith E.
/ TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
/ FILE REFERENCE: 5800-119
/ CURRENT APPLICATION NUMBER: US/09/833,381
/ CURRENT FILING DATE: 2001-04-11
/ PRIOR APPLICATION NUMBER: 09/516,448
/ PRIOR FILING DATE: 2000-02-29
/ NUMBER OF SEQ ID NOS: 2050
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 252
/ LENGTH: 422
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-833-381-252
```

```
Query Match          2.9%; Score 145.2; DB 10; Length 422;
Best Local Similarity 87.4%; Pred. No. 2.5e-35;
Matches 159; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
```

```
QY 209 TGCGTGTATCATGAACTACTGACAGAAATCCAGATGCTGTGGACGCTTATTGTTATAC 268
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150 TGACTGTATCATGAACTACTGACAGAAATCCAGATGCTGTGGAGGCCCTGTGTTATAC 209

QY 269 GAGGATCCCGGCTGTCAAGTGGAGTCTCAACCTGACGCAATGCTTCAGACCCAAAG 328
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 210 GATGATCCCAAGTCAAGTGGAGTCAACCTGACGCAATGCTTCAGACCCAAAG 269

QY 329 GACTGCGGCTGCGCTCCGACTGTATCCCGGTTCCAAAGCTGAGAGCTTCCTCCGACA 388
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 270 GACTGCGAGTGTGCTGTGACTGTATCCCGGTTCCAAAGCTGAGAGATTCCTCCAAAC 329

QY 389 AG 390
      ||
Db 330 AG 331
```

```
RESULT 14
US-09-960-352-14383
/ Sequence 14383, Application US/09960352
/ Patent No. US20020137139A1
/ GENERAL INFORMATION:
/ APPLICANT: Warren, Wesley C.
/ APPLICANT: Tao, Mengding
/ APPLICANT: Byatt, John C.
/ APPLICANT: Mathialagan, Nagappan
/ TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
/ FILE REFERENCE: 16511.006/37-21(10298)C
/ CURRENT APPLICATION NUMBER: US/09/960,352
/ CURRENT FILING DATE: 2001-09-24
/ NUMBER OF SEQ ID NOS: 15112
/ SEQ ID NO 14383
/ LENGTH: 234
/ TYPE: DNA
/ ORGANISM: Bos taurus
/ OTHER INFORMATION: Clone ID: 61-LIB34-051-Q1-E1-H2
US-09-960-352-14383
```

```
Query Match          2.8%; Score 139.6; DB 10; Length 234;
Best Local Similarity 74.8%; Pred. No. 9.5e-34;
Matches 175; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
```

```
QY 4202 GACTGCTGCTCAGCTGTGAAAGATCTCAAGGCTTCATCTTACAGGTATCTCTGG 4261
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 GACTGCTGCTCAGCTGTGAAAGATCTCAAGGCTTCATCTTACAGGTATCTCTGG 60

QY 4262 TGCAACCAAGATGAACCTGCAATCTCATGTTCAAGAAATGAAGTGTAGGCTGTT 4321
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 TGCAACCAAGATGAAGTGTGCAAGATGTTCAAGAAATGAAGTGTAGGCTGTT 120

QY 4322 CTTGGAGCCCAACAGACAGATATTTGCTTGAAGCTTACAGAGGCTGCTCCATACAC 4381
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 CCGGAGGCTCTCTCAGAGGAGCATTTGCTTACACTGACAGACCTGCTCATATCACC 180

QY 4382 TGACAAATTAATGCAAGCTTGTGCAATCCCAAGCTTACATGCTTACACCCGAC 4435
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 GAAAGAGTAAATCCAGCTTGTGCAACCCCAAGCTTACATGCTTGTGACCCG 234
```

```
RESULT 15
US-09-960-352-11543
/ Sequence 11543, Application US/09960352
/ Patent No. US20020137139A1
/ GENERAL INFORMATION:
/ APPLICANT: Warren, Wesley C.
/ APPLICANT: Tao, Mengding
/ APPLICANT: Byatt, John C.
/ APPLICANT: Mathialagan, Nagappan
/ TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
/ FILE REFERENCE: 16511.006/37-21(10298)C
/ CURRENT APPLICATION NUMBER: US/09/960,352
```


Gencore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 05:15:44 ; Search time 5999.5 Seconds
(without alignments)
13497.364 Million cell updates/sec

Title: US-09-923-515-3_COPY_8939_13938

Perfect score: 5000
Sequence: 1 ctcgcactcttaccgcggtt.....ctgcactatttgattga 5000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pla:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	606	12.1	856	9	AL579900 AL579900
2	601.4	12.0	766	13	BI759134 603042546
3	568	11.4	568	12	BF970104 602272382
4	546.8	10.9	766	10	AV700701 AV700701
5	502	10.0	859	12	BG402207 602465764
6	461.4	9.2	788	10	AV695776 AV695776

C	7	441.4	8.8	622	9	AI061613	AI061613 HA0380 Hu
C	8	396	7.9	762	12	BG428018	BG428018 602501424
C	9	395.8	7.9	750	10	AV695796	AV695796 AV695796
C	10	392	7.8	593	10	BE348267	BE348267 hw21e11.x
C	11	387.4	7.7	701	13	BI219735	BI219735 602936156
C	12	383.4	7.7	611	10	AV662084	AV662084 AV662084
C	13	370	7.4	481	14	H60805	H60805 yr45909.r1
C	14	368.6	7.4	563	10	AM721976	AM721976 xs12d10.x
C	15	365	7.3	758	12	BG616686	BG616686 602614893
C	16	364.2	7.3	668	12	BG427320	BG427320 602494402
C	17	361.6	7.2	594	10	AV661991	AV661991 AV661991
C	18	359.2	7.2	558	10	AM104579	AM104579 xdf6f03.x
C	19	357	7.1	553	10	AT633979	AT633979 wa32c10.x
C	20	355.4	7.1	466	9	AT948806	AT948806 wg22a02.x
C	21	350	7.0	504	14	N77239	N77239 yv44d01.t1
C	22	348.2	7.0	541	9	AA722885	AA722885 z989906.s
C	23	348.2	7.0	740	10	AV653775	AV653775 AV653775
C	24	346.8	6.9	636	10	AV645329	AV645329 AV645329
C	25	346.4	6.9	515	9	AT351543	AT351543 qf06e10.x
C	26	345.4	6.9	568	14	T62069	T62069 ycf66e05.s1
C	27	341.2	6.8	462	9	AT768449	AT768449 wh22a07.x
C	28	340.8	6.8	545	10	AV720081	AV720081 AV720081
C	29	340.8	6.8	621	9	AT377474	AT377474 tc37a03.x
C	30	340.6	6.8	464	10	BE326689	BE326689 hr63d09.x
C	31	340.6	6.8	595	10	AM950594	AM950594 EST362664
C	32	339	6.8	457	9	AT631756	AT631756 wa35901.x
C	33	334.6	6.7	462	9	AA995929	AA995929 ou93h04.s
C	34	333.8	6.7	444	9	AT640315	AT640315 wa16908.x
C	35	332.6	6.7	483	10	AV720676	AV720676 AV720676
C	36	328.6	6.6	607	10	AV662061	AV662061 AV662061
C	37	325.8	6.5	466	9	AL531542	AL531542 AL531542
C	38	325.6	6.5	755	13	BI145879	BI145879 602911179
C	39	324.8	6.5	614	10	AV693554	AV693554 AV693554
C	40	321.2	6.4	586	9	AA106793	AA106793 mm20e04.t
C	41	319.6	6.4	424	14	H73861	H73861 ysl2a02.t1
C	42	319.2	6.4	687	10	AV650465	AV650465 AV650465
C	43	318.2	6.4	694	9	AA032930	AA032930 ml22d06.r
C	44	318	6.4	564	10	BE65711	BE65711 154843 MA
C	45	318	6.4	668	14	BQ598303	BQ598303 MI-P-E3-a

ALIGNMENTS

RESULT 1
AL579900/c
LOCUS AL579900 LTI_NFL008.TC2 Homo sapiens cDNA clone CSDD002YJ21 3
DEFINITION prime, mRNA sequence.
ACCESSION AL579900
VERSION AL579900.1 GI:12945394
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 856)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source location/Qualifiers
1. 856
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDD002YJ21"
/clone_lib="LTI_NFL008.TC2"
/sex="male"
/tissue_type="T cells from T cell leukemia"
/note="Vector: pcMVSPORT 6; Site_1: NotI; 1st strand cDNA

Chinese National Human Genome Center at Shanghai
331 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanz@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES

SOURCE

1. 766
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKAEH09"
/clone_lib="GKC"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 199 a 204 c 166 g 193 t 4 others
ORIGIN

Query Match 10.3%; Score 546.8; DB 10; Length 766;
Best Local Similarity 84.0%; Pred. No. 3,8e-152;
Matches 657; Conservative 0; Mismatches 81; Indels 44; Gaps 2;

Oy 4214 CTGCTTGAAGAGTCTCAAGGCTTATCTCAAGAGTATCTGGGTCACACCAAGA 4273
|||||
Db 765 CTGCTTGAAGAGTCTCAAGGCTTATCTCAAGAGTATCTGGGTCACACCAAGA 706
Oy 4274 AGTGAACCTCGAATCTCATCTTTCAGAAATAGAAGTGTCTAGGCTCTTCTGGAGCCAC 4333
|||||
Db 705 AGTGAATCTCGAATCTCATCTTTCAGAAATAGAAGTGTCTAGGCTCTTCTGGAGCCAC 646
Oy 4334 ACAAGCAGATTTTCCTTCTGTAAGTACAGAGCCCTGCGCTATCATCTGACAAATAT 4393
|||||
Db 645 ACAAGCAGATTTTCCTTCTGTAAGTACAGAGCCCTGCGCTATCATCTGACAAATAT 586
Oy 4394 GCCAGCTTGTGCGATCCCGACAGCTACATGTCACCCGCGACGATGATGATCAG 4453
|||||
Db 585 GCCAGCTTGTGCGATCCCGACAGCTACATGTCACCCGCGACGATGATGATCAG 526
Oy 4454 TGGCTGGGAGAAACCCAGGATCTTGGGATCGGCTTCTCAAGAACCCAGCTCCT 4513
|||||
Db 525 TGGCTGGGAGAAACCCAGGATCTTGGGATCGGCTTCTCAAGAACCCAGCTCCT 466
Oy 4514 TGTATTGAGAAATGATGTCATCTCAATCAATGATGATGATGATGATGATGATGAT 4548
|||||
Db 465 TGTATTGAGAAATGATGTCATCTCAATCAATGATGATGATGATGATGATGATGAT 406
Oy 4549 - -ATATTGTCCTAGCATTTGGCCAGAGCAGTGTGCGAGGATGAGAGTGTGAGG 4606
|||||
Db 405 CGAAGCTGTCTGCGATTTGGCCAGAGCAGTGTGCGAGGATGAGAGTGTGAGG 346
Oy 4607 GCGCTGCTTGTCTGCGAGAGCAATATCTTTCAGAGGATCACTTCTGGGCTCT 4666
|||||
Db 345 GCGCTGCTTGTCTGCGAGAGCAATATCTTTCAGAGGATCACTTCTGGGCTCT 286
Oy 4667 TGGCTGTGACGCCCCCAATAGCTGTGTATGCTGCTGTTCAAGGTTTGTACTTG 4726
|||||
Db 285 TGGCTGTGACGCCCCCAATAGCTGTGTATGCTGCTGTTCAAGGTTTGTACTTG 226
Oy 4727 GATTGAGGAGTATGATGAATTAATTAATGAGAGGAGAGCAGAGAGCAATCACTA 4786
|||||
Db 225 GATTGAGGAGTATGATGAATTAATTAATGAGAGGAGAGCAGAGCAATCACTA 166
Oy 4787 CTTAAGAGCTGAAGCTGGGTAGAGTTAGCATGCTGGAATTAATTAATCAAGCAATCA 4846
|||||
Db 165 CCTGAGAGCTGGAGCTGGGTAGAGTTAGCATGCTGGAATTAATTAATCAAGCAATCA 106
Oy 4847 CGAAGAGCTGTGCGAGCTAGCAGCTATGCAAACTTGGCATTTTGTATTTTGG 4906
|||||
Db 105 CGAAGAGCTGTGCGAGCTAGCAGCTAGCAGCAACTTGGCATTTTGTATTTTGG 49

Oy 4907 TATAGCTTTTAACTGACAGCAATTCCTTATTAAGGTCTCACTATGACATTT 4966
|||||
Db 48 -----TCTGACTGCTGATCTGTAGTAAAGTGACATACCTATGACATTT 3

Oy 4967 GT 4968
||
Db 2 GT 1

RESULT 5
BG402207 859 bp mRNA linear EST 12-MAR-2001
LOCUS 602465764F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4594023 5',
DEFINITION mRNA sequence.

ACCESSION BG402207
VERSION BG402207.1 GI:13295655

KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (bases 1 to 859)
NIH-MGC <http://imgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs@email.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LMC134 row: 0 column: 16
High quality sequence stop: 629.
Location/Qualifiers

FEATURES

SOURCE

1. 859
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4594023"
/clone_lib="NIH_MGC_75"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: kidney; Vector: pNR-LIB (Clontech); Site_1:
SfiI (ggcgccctggc); Site_2: SfiI (ggccatcagcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCATTAAGGCC-3' and 3' adaptor sequence:
5'-ATTCTGAGGCGGAGCGGCGCGCATG-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.65
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH-MGC Library."

BASE COUNT

183 a 233 c 242 g 201 t

ORIGIN

Query Match 10.0%; Score 502; DB 12; Length 859;
Best Local Similarity 87.8%; Pred. No. 1.2e-138;
Matches 571; Conservative 0; Mismatches 75; Indels 4; Gaps 2;

Oy 3899 TCCAGGAGCAATTAATGAGGAGCTGTGAAAAAATTAATGCGCTTAACCTGATGTGA 3958
|||||
Db 1 TCCAGGAGCAATTAATGAGGAGGAGGCTGTGAAAAAATTAATGCGCTTAACCTGATGTGA 60
Oy 3959 CATCAATGAGTCCCTGGTGTCTACATGATCCAGAAAACTTTTGTACTACTGTGATAT 4018
|||||
Db 61 TGTAGAGTGTCCCTGGTGTCTACATGATCCAGAAAACTTTTGTACTACTGTGATAT 120
Oy 4019 CCTCTGTGTCATCTCTTCATTTGATTTGTGGAAGCCCTCAAGTGGAGCCGAAAGATG 4078
|||||
Db 121 CCTCTGTGTCATCTCTTCATTTGATTTGTGGAAGCCCTCAAGTGGAGCCGAAAGATG 180
Oy 4079 TCTTGAAGCATTTGTAGGGGGGTGTGGCCACCAACCAATCTTGGCCCTGGCAAGTCA 4138

```
Db 181 TCCGTGAAAGGGTTGTAGGGGGGTGTGGCCCAACCCATCTCTGGCCCTGGCAAGTCAG 240
QY 4139 TCTCAGACAAAGGTTTGAAGACACTTCTGTGAGGACACCTTAATATCCAGAGTGGT 4198
Db 241 TCTTAGACAAAGTTTGAATGACTTCTGTGAGGACACCTTGAATATCCAGAGTGGT 300
QY 4199 GCTGATGCTGCTCATGCTGCTTGAAGAGTCTCAAGGCTTCACTTAAAGTCTATCTT 4258
Db 301 GTTGACTGCTGGCCACTGCTTGGAGAGTCCCAAGGCTTATCCTCAAGGTCATCTCT 360
QY 4259 GGGTGCACACCAAGTGAACCTCGAATCTCATCTTCAAGAAATGAAGTCTTAGGCT 4318
Db 361 GGGTGCACACCAAGTGAATCTCGAACGCACTGTTCAAGAAATGAAGTCTTAGGCT 420
QY 4319 GTTCTTGGAGCCCAACACAGCATATTTGCTTGTAAAGCTAGACAGGCTCCGCTCAT 4378
Db 421 GTTCTTGGAGCCCAACAGAAATATTTGCTTGTAAAGCTAGACAGTCTCCGCTCAT 480
QY 4379 CACTGACAAAGTAAATGACACTTGTCTGCATCCCACTACATGTCACCGCCAGAC 4438
Db 481 CACTGACAAAGTAAATGACACTTGTCTGCATCCCACTACATGTCGCTGACCGGAC 540
QY 4439 TGAATGTCATCATCT--GGCTGGGAGAAACCCAAAGTACCTTTGGGA--CTGGCCTTC 4494
Db 541 CGAATGTCATCATCTGCTTGGGAGAAACCCAAAGTACCTTTGGGAGCTGGCCTTTC 600
QY 4495 TCAAGGAGCCGCTCCTCTGTTATGAGATGAATGATGTCATCTACTAT 4544
Db 601 TCAAGGAGCCGCTCCTCTGTTATGAGATGAATGATGTCATCTACTAT 650

RESULT 6
AV695776 788 bp mRNA linear EST 16-JAN-2002
LOCUS AV695776 GKC Homo sapiens cDNA clone GKCXB08 5', mRNA sequence.
DEFINITION AV695776
ACCESSION AV695776.1 GI:10297639
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 788)
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE 21625106
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzgchgc.sh.cn
FEATURES
source
location/Qualifiers
1..788
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GKCXB08"
/clone_1lb="GKC"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 196 a 195 c 227 g 167 t 3 others
```

```
ORIGIN
Query Match 9.28; Score 461.4; DB 10; Length 788;
Best Local Similarity 85.3%; Pred. No. 1,8e-126;
Matches 538; Conservative 0; Mismatches 86; Indels 5; Gaps 2;

QY 3584 CCCGAAAACATACCCAAATGATGCTGACATGATCTACTGACAGAAATCCAGATCCGA 3643
Db 2 CCCGAAAACATACCCAAATGCTGGCTGACATGATCTACTGACAGAAATCCAGATCCGA 61
QY 3644 TACAGG--CCCTTGGTGTTTTACATGAGACCCCAACATCAGTGGAGTACTGCAACCTGA 3702
Db 62 TAAAGGCCCTTGGTGTTTTACCAAGACCCCAACGCTCAGAGTGGAGTACTGCAACCTGA 121
QY 3703 CGCATGCTCAGACAGAAAGGACGTGGTGGCTCTCCGACGTGATCAGAGTTCGAC 3762
Db 122 AAAATGCTCAGAGACAGAGCGAGTGTGAGACCTCCGCTGTCTGCTTCAG 181
QY 3763 GCTTAGGCTCTCTTCTGAAACAGACTGATGTTTGGAAATGGGAAAGATATACCGGGCA 3822
Db 182 ATGTAGAGACTCTCTCCGAAGAGACTGTATGTTGGAAATGGGAAAGATATACCGGGCA 241
QY 3823 AGAAGGCAACCACTGTACTGAGACGCCATGACAGAAATGGGCTCCCAAGACCCATA 3882
Db 242 AGAGGCGACCACTGTACTGAGACGCCATGACAGAAATGGGCTCCCAAGACCCATA 301
QY 3883 GACACAGCAGTTCATTCACAGGACAAATTAATGGCAGGCTGGAATAAATTTACTGCC 3942
Db 302 GACACAGCAGTTCATTCACAGGACAAATTCACAGGCGGGGCTGGAATAAATTTACTGCC 361
QY 3943 GTAACCTGATGATGATCATCAATGATGCTGCTACACATGATATCCAGAAACTTT 4002
Db 362 GTAACCTGATGATGATGATGATGATGCTGCTACACAGCAAAATCCAG--CTTT 417
QY 4003 TTGACTAGTGAATATCCCTCTGTCATCTCTGATTTGATTTGGAGAGCTTCAAG 4062
Db 418 ACGACTAGTGAATATCCCTCAATGATGAGGCGCTTCAATTTGATTTGGAGAGCTTCAAG 477
QY 4063 TGGAGCCAGAAATGCTCTGGAAGCATTTAGAGGGGTGTGTGGCCACCCATTCCT 4122
Db 478 TGGAGCCAGAAATGCTCTGGAAGCATTTAGAGGGGTGTGTGGAGCCACCCATTCCT 537
QY 4123 GGCCCTGGCAAGTCAAGTCTCAAGCAAGTTTGAAGCACTTCTGTGGAGGCACTTAA 4182
Db 538 GGCCCTGGCAAGTCAAGTCTTGAACAAAGTTTGAAGCACTTCTGTGGAGGCACTTGAAT 597
QY 4183 TATCCCAAGTGGGTGCTGACTGCTGCT 4211
Db 598 ATCCCAAGTGGGTGAGTGGACTGTGCCACT 626

RESULT 7
A1061613 622 bp mRNA linear EST 11-NOV-1999
LOCUS A1061613/c
DEFINITION HA0380 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
sequence.
ACCESSION A1061613
VERSION A1061613.1 GI:6358910
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 622)
AUTHORS Yu,Y., Zhang,C., Luo,L., Ouyang,S., Li,W., Wu,J., Zhou,S., Liu,M.,
and He,F.
TITLE Expression profile analysis of a human fetal liver cDNA library
JOURNAL Unpublished (1998)
COMMENT Contact: Yongtao Yu
Department of Hematology
Beijing Institute of Radiation Medicine
27 Taiping Road, Beijing 100850, P.R.China
Tel: 0086-10-68159479
```


4800 ACCTGGTAAAGATTATTCATCTGTGAATATATAGACAGCAATCAACGAGACAC-TGT 4858
|||||
481 ACCTGGTAAAGATTATTCATCTGTGAATATATAGACAGCAATCAACGAGACAC-TGT 540
|||||
4859 TCCGAGTACACAGCTATCCCAACCTTGGCATTTTGGTATTTTGGTATTA 4918
|||||
541 CCCGAGTACACAGCTATCCCAACCTTGGCATTTTGGTATTTTGGTATTTT 600
|||||
4919 AGGTCTGACTGACAAATCTGTATTAAGTGTCTATA 4954
|||||
601 GTAGTAGGTGACATAGTATGATCTGTATTA 636
|||||
RESULT 9
AV695796 750 bp mRNA linear EST 16-JAN-2002
LOCUS AV695796 GK Homo sapiens cDNA clone GCGXB1 5', mRNA sequence.
DEFINITION AV695796
ACCESSION AV695796
VERSION AV695796.1 GI:10297659
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 750)
Xiao,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Shen,K., Lu,G., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
JOURNAL MEDLINE
COMMENT Contact: Zenguan Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhanjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzge@hgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
1..750
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKGBX11"
/clone_lib="GKC"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI;"
BASE COUNT 179 a 184 c 229 g 157 t 1 others
ORIGIN
Query Match 7.9%; Score 395.8; DB 10; Length 750;
Best Local Similarity 80.3%; Pred. No. 8.7e-107;
Matches 551; Conservative 0; Mismatches 73; Indels 62; Gaps 5;
QY 3584 CCCAGAAACTACCAATGATGCTTCAGCAATGAGTACAGGATCCAGATGCCGA 3643
DB 4 CCCAGAAACTACCAATGATGCTTCAGCAATGAGTACAGGATCCAGATGCCGA 63
QY 3644 TACAGG-CCCTTGTGTATTACATGACCCAGCATCAGTGGAGTACTGCAACCTGA 3702
DB 64 TAAAGGCCCTGGTGTATTACACACAGCCAGCGTCAGGGAGTACGACACCTGA 123
QY 3703 CCGGATGCTCAACACAGAGGAGTGTGCTCCTCCGACGTCTATCCAGGTCCAA 3762
DB 124 AAAAATGCTCAGCAAGAGGAGTGTGAGCACTCCGCTGTGCTCTGCTCCAG 183
QY 3763 GCCTAGGGCTCTCTCTGACAAAGACTGTATGTTGGGAATGGAAGA-TACCGGGGC 3821

DB 184 ATGTAGAGACTCTCTCCAGAAAGACTGTATGTTGGGAATGGGAAGGATACCGAGGC 243
QY 3822 AAGAAGGCAACACCTGTACTGGGACGCATGCCAGGAATGGGCTGCCAGGACCCCAT 3881
DB 244 AAGAGGCGACACCTGTACTGGGACGCATGCCAGGAATGGGCTGCCAGGACCCCAT 303
QY 3882 AGACACAGCAGCTTATTCAGGACAAATTAATGGGCAAGGTCTGGAAAAATTAATGTC 3941
DB 304 AGACACAGCAGCTTATTCAGGACAAATTAATGGGCAAGGTCTGGAAAAATTAATGTC 359
QY 3942 CGTAACCTGATGGGACATCAATGCTCCCTGGTGTCTACAAATTAATTAAGAAACT 4001
DB 360 CGTAACCTGATGGGACATCAATGCTCCCTGGTGTCTACAAATTAATTAAGAAACT 411
QY 4002 TTGTACTGTGATATCCCTCTCTGTGCATCCCTTCATTTGATGTGGGAACCTCAA 4061
DB 412 -----GTGGGAAGCTCAA 425
QY 4062 GTGGAGCCGAATAATGTCTTGGAGCATGTAGGGGGGTGTGGCCACCCACATTC 4121
DB 426 GTGGAGCCGAATAATGTCTTGGAGCATGTAGGGGGGTGTGGAGCCACCCACATTC 485
QY 4122 TGGCCCTGGCAAGTCACTCAGAACAGTTTGGAAAGCACTTGTGGAGCAGCTTA 4181
DB 486 TGGCCCTGGCAAGTCACTCAGAACAGTTTGGAAAGCACTTGTGGAGCAGCTTG 545
QY 4182 ATATCCCAAGAGTGGTGTCTACTGCTGTCTGAGAGAGTCTCAAGGCTTCA 4241
DB 546 ATATCCCAAGAGTGGTGTCTACTGCTGTCTGAGAGAGTCTCAAGGCTTCA 603
QY 4242 TCCATCAAGTCACTGCTGGTGCACA 4267
DB 604 ATCTACAGTCACTGCTGGTGCACA 629
RESULT 10
BE348267 593 bp mRNA linear EST 18-JUL-2000
LOCUS hw21ell.x1 NCI CGAP Kid1 Homo sapiens cDNA clone IMAGE:318356 3'
DEFINITION similar to gb:X05199 PLASMINOGEN PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION BE348267
VERSION BE348267.1 GI:9260120
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 593)
AUTHORS NCI CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
Unpublished (1997)
JOURNAL COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgap@rt.ema1.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL, send email to:
info@image.lnl.gov
Seq primer: -400p from Glbco
High quality sequence stop: 425.
FEATURES
source
1..593
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:318356"
/clone_lib="NCI CGAP Kid1"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: pTY73D-Pac (Pharmacia) with

ACCESSION BG616686
VERSION BG616686.1 GI:13668057
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS 1 (bases 1 to 758)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rsb@biml.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LUCM1595 row: j column: 24
High quality sequence stop: 553.
Location/Qualifiers
1..758
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4733687"
/clone_lib="NIH-MGC_76"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: liver; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgctggcgc); Site_2: SfiI (ggccatgatggc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATATGAGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.85
kb (range 1.0-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH-MGC Library."

BASE COUNT 242 a 150 c 206 g 160 t
ORIGIN

Query Match 7.3%; Score 365; DB 12; Length 758;
Best Local Similarity 82.7%; Pred. No. 1.6e-97;
Matches 459; Conservative 0; Mismatches 55; Indels 41; Gaps 2;

QY 4462 GAGAAACCAAGTACTTTGGGACTGGCTTCTCAAGAGCCAGCTCTTGTATATG 4521
|||||
Db 1 GAGAAACCAAGTACTTTGGGACTGGCTTCTCAAGAGCCAGCTCTTGTATATG 60

QY 4522 AGAATGAAGTGTCAATCTCTTAAGT-----ATATTT 4554
|||||
Db 61 AGAATGAAGTGTCAATCTCTTAAGT-----ATATTT 120

QY 4555 GTGCTGAGCATTTGGCCAGGCTGACAGTGGAGGGGCTCTGG 4614
|||||
Db 121 GTGCTGAGCATTTGGCCAGGCTGACAGTGGAGGGGCTCTGG 180

QY 4615 TTTGCTTCGAGAAGCAAAATCAATTTTACAAGGAGTCACTTTGGGGTCTTGGCTGTG 4674
|||||
Db 181 TTTGCTTCGAGAAGCAAAATCAATTTTACAAGGAGTCACTTTGGGGTCTTGGCTGTG 240

QY 4675 CACGCCCAATAAGCTGTGTCTATGCTCGTGTTCAGAGTTTGTACTTGGATTGAGG 4734
|||||
Db 241 CACGCCCAATAAGCTGTGTCTATGCTCGTGTTCAGAGTTTGTACTTGGATTGAGG 300

QY 4735 GAATGATGAGAAATATTAATTTGACGGGAGACAGAGTGAACATCAACCTTATGAAG 4794
|||||
Db 301 GAATGATGAGAAATATTAATTTGACGGGAGACAGAGTGAACATCAACCTTATGAAG 360

QY 4795 CTGAAACGTGGGTAGGATTTAGCATGCTGGAATATATATAGACAGCAATCAAGAGACA 4854
|||||
Db 361 CTGAAACGTGGGTAGGATTTAGCATGCTGGAATATATATAGACAGCAATCAAGAGACA 420

QY 4855 CTGTCCAGCTACACGCTATGCAAAACCTTGGCATTTTGGTATTTGTGATTAAGCT 4914
|||||
Db 421 CTGTCCAGCTACACGCTATGCAAAACCTTGGCATTTTGGTATTTGTGATTAAGCT 480

QY 4915 TTTAAGTCTGAGTACAAATTCGTATTAAGGTGTCAATAGCTATGACATTTGTAAAAA 4974
|||||
Db 481 TTTAAGTCTGAGTACAAATTCGTATTAAGGTGTCAATAGCTATGACATTTGTAAAAA 526

QY 4975 TAAACTCTGCACCTTA 4989
|||||
Db 527 TAAACTCTGCACCTTA 541

Search completed: March 6, 2003, 15:20:05
Job time : 6051.5 secs

